



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Distributive ignorance inferences with wonder and believe

Citation for published version:

Cremers, A, Roelofsen, L & Uegaki, W 2019, 'Distributive ignorance inferences with wonder and believe', *Semantics and Pragmatics*, vol. 12, no. 5, 5:EA, pp. 1-60. <https://doi.org/10.3765/sp.12.5>

Digital Object Identifier (DOI):

[10.3765/sp.12.5](https://doi.org/10.3765/sp.12.5)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Semantics and Pragmatics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



This is an EARLY ACCESS version of

Cremers, Alexandre, Floris Roelofsen & Wataru Uegaki. 2019. Distributive ignorance inferences with *wonder* and *believe*. *Semantics and Pragmatics* 12(5). 1–60. <https://doi.org/10.3765/sp.12.5>.

This version will be replaced with the final typeset version in due course.
Note that page numbers will change, so cite with caution.

Distributive ignorance inferences with *wonder* and *believe**

Alexandre Cremers
ILLC, University of Amsterdam

Floris Roelofsen
ILLC, University of Amsterdam

Wataru Uegaki
LUCL, Leiden University

Abstract A sentence like *Mary wonders whether Ann, Bill or Carol broke the vase* implies that Mary still consider all disjuncts possible. This inference has been referred to as a *distributive ignorance inference* (Roelofsen & Uegaki 2016). We present two experiments examining the distributive ignorance inferences triggered by two verbs, *wonder* and *believe*, with different types of complements and different types of quantificational subjects.

The results of these experiments show that the distributive ignorance inferences triggered by the two verbs pattern very much alike. We argue that the data are best explained by an account that involves a strengthening mechanism which is sensitive to the syntactic structure of the complement of the verbs involved and optionally applies locally, as part of the semantic composition process.

Keywords: ignorance inferences, *wonder*, *believe*, local strengthening

1 Introduction

Consider the following sentence:

- (1) The detective wonders whether Ann, Bill, or Carol did it.

This sentence implies that the detective doesn't know yet whether Ann did it, that he doesn't know yet whether Bill did it, and that he doesn't know yet whether Carol did it. Roelofsen & Uegaki (2016) call this *distributive ignorance*. Note that distributive ignorance is a stronger form of ignorance than merely not knowing the answer to the embedded question in (1), i.e., the question whether Ann, Bill, or Carol did it.

* We are grateful to Kyle Blumberg, Ivano Ciardelli, Jakub Dotlačil, Clemens Mayr, Yael Sharvit and Yimei Xiang for discussion. We would also like to thank our *S&P* editor Kristen Syrett, Aaron Steven White, and two anonymous reviewers for very helpful comments on earlier versions of the paper. Finally, we gratefully acknowledge financial support from the Netherlands Organisation for Scientific Research (NWO) and the European Research Council (ERC, grant agreement number 680220).

If the detective already knows that Carol didn't do it, but still wonders whether it was Ann or Bill, then he doesn't yet know the answer to the question whether Ann, Bill, or Carol did it. Thus, he is still ignorant to some extent. Distributive ignorance, however, requires more than this: the detective should still consider all three options possible.

Roelofsen & Uegaki (2016) show that the distributive ignorance inferences that sentences like (1) give rise to are not predicted by previous work on the semantics of *wonder* (Ciardelli & Roelofsen 2015, Uegaki 2015), even if pragmatic strengthening is taken into account. Based on various further empirical observations, they develop a refined semantic entry for *wonder*.

It seems, however, that distributive ignorance inferences (henceforth, DIIs) may well constitute a broader phenomenon, not specific to *wonder*. In particular, epistemic predicates like *believe* seem to give rise to similar inferences. To see this, consider the sentence in (2):

- (2) The detective believes that Ann, Bill, or Carol did it.

Just like (1), this sentence also seems to imply that the detective still considers all of Ann, Bill, and Carol possible culprits.

The goal of the present paper is to investigate experimentally whether the ignorance inferences triggered by *wonder* and *believe* indeed call for a unified account, and if so, what the main empirical desiderata for such an account are.¹ More specifically, we report the results of two experiments, each of which is designed to address a basic question about the nature of DIIs triggered by *wonder* and *believe*. The first experiment addresses the question whether DIIs are *structure-sensitive*. More specifically, we examine the extent to which DIIs persist when *wonder* and *believe* take different kinds of complements. For instance, sentences like (1) and (2), which involve disjunctive complements, are compared with sentences like (3) and (4) below, which involve a wh-phrase and existential quantification, respectively.

- (3) The detective wonders which of the girls did it.

- (4) The detective believes that one of the girls did it.

Our second experiment addresses another basic issue, namely whether DIIs are computed locally, as part of the semantic composition process, or rather result from global pragmatic reasoning. More specifically, we examine what kind of ignorance inferences arise if *wonder* and *believe* take different kinds of quantificational sub-

¹ Besides *wonder* and *believe*, many other predicates that have some inquisitive or epistemic meaning component seem to trigger DIIs as well (e.g., *be curious*, *investigate*, *suspect*, *be certain*, *hope*). We restrict ourselves here to *wonder* and *believe*, leaving a detailed investigation of this broader range of predicates for future work.

jects (upward monotonic, downward monotonic, or non-monotonic) rather than a referential expression like *the detective*. For instance, we consider sentences like (5) and (6) below. Whether such sentences are judged true or false in certain scenarios crucially depends on whether DIIs are computed locally or globally.²

- (5) Every detective wonders whether Ann, Bill, or Carol did it.
- (6) No detective believes that Ann, Bill, or Carol did it.

The overall results of the two experiments show that DIIs triggered by *wonder* and *believe* behave very similarly. This means that the results do not provide any grounds to reject the hypothesis that DIIs triggered by *wonder* and *believe* arise from the same mechanism. Of course we cannot exclude the possibility that future experimental findings will refute this hypothesis. However, in view of the experimental data gathered here, the most parsimonious account is one that derives DIIs triggered by *wonder* and *believe* in a unified way. Moreover, the results also provide some important clues as to what such a unified account should look like. In particular, the patterns we find suggest that DIIs result from a strengthening mechanism which is sensitive to the syntactic structure of the complement of the verbs involved (cf., Katzir 2007) and optionally applies locally, as part of the semantic composition process (cf., Chierchia et al. 2012). We spell out a concrete account of DIIs that is compatible with these findings.

The paper is organized as follows. In Section 2 we introduce our baseline lexical entries for *believe* and *wonder* and show that these entries by themselves do not account for DIIs. In Section 3 we outline various possible refinements of this baseline account, and discuss how these may be teased apart experimentally. In Sections 4 we present the results of our first experiment, involving different kinds of complements. In Section 5 we turn to our second experiment, involving different kinds of quantificational subjects. Finally, Section 6 provides a general discussion of the results and explicates the theoretical approach that they support in further detail. Section 7 concludes.

2 A baseline account of *wonder* and *believe*, and its limitations

For concreteness, we will fix a specific baseline account of *wonder* and *believe*, and the complements that they take. For *wonder* we will adopt the analysis proposed in Ciardelli & Roelofsen (2015), which is formulated in *inquisitive semantics* (Ciardelli

² As we will see, our experimental results go beyond what could have been established by eliciting judgments from a small sample of informants. The quantitative differences between some of the target conditions are relatively subtle, and some conditions show considerable variation between speakers.

et al. 2013, 2018).³ For *believe* we will adopt the canonical lexical entry, rooted in epistemic logic (Hintikka 1962). For uniformity we will formulate this entry in inquisitive semantics as well, though nothing in our discussion will hinge on this. We will briefly review the relevant notions from inquisitive semantics (§2.1) and the semantic analysis of declarative and interrogative complements we assume (§2.2). Then we will spell out our baseline account of *wonder* and *believe* (§2.3-2.4), and finally we will show that this account by itself falls short of deriving DIIs for sentences like (1) and (2) (§2.5).

2.1 Inquisitive semantics background

In inquisitive semantics, declarative and interrogative clauses are taken to have the same kind of semantic value, namely a set of propositions. The conceptual motivation behind this uniform notion of sentence meaning is as follows. While traditionally the semantic value of a sentence φ is intended to capture just the information conveyed by φ , in inquisitive semantics it is intended to capture the *issue* expressed by φ as well. To achieve this, the semantic value of φ , $\llbracket \varphi \rrbracket$, is construed as the set of those propositions that:

- (i) resolve the issue that φ expresses (if any) and
- (ii) do not contain any possible worlds that are ruled out by the information that φ conveys (if any).

For instance, the semantic value of the declarative sentence *Ann left* is the set of propositions consisting exclusively of worlds in which Ann left.

$$(7) \quad \llbracket \text{Ann left} \rrbracket = \{p \mid \forall w \in p : \text{Ann left in } w\}$$

Let us check that these are indeed the propositions that satisfy conditions (i) and (ii) above. First, since the sentence does not express any (non-trivial) issue, condition (i) is automatically satisfied. On the other hand, the sentence does convey the information that Ann left, so condition (ii) is only satisfied by those propositions that do not contain any worlds in which Ann didn't leave. This is exactly the set of propositions in (7).

Similarly, the semantic value of the interrogative sentence *Did Ann leave?* is the set of propositions which either consist exclusively of worlds in which Ann left, or exclusively of worlds in which Ann didn't leave.

³ See Uegaki (2015) for a closely related analysis, and Theiler et al. (2017a) for comparison. For earlier informal discussions of the semantics of *wonder*, see Karttunen (1977) and Guerzoni & Sharvit (2007).

$$(8) \quad \llbracket \text{Did Ann leave?} \rrbracket \\ = \{p \mid \forall w \in p : \text{Ann left in } w\} \cup \{p \mid \forall w \in p : \text{Ann didn't leave in } w\}$$

In this case, since the sentence does not convey any (non-trivial) information, condition (ii) is automatically satisfied. On the other hand, condition (i) is only satisfied by propositions which resolve the issue whether Ann left, i.e., ones which either establish that Ann did leave or that she didn't. This is exactly the set of propositions in (8).

Downward-closure and alternatives The set of propositions associated with a sentence φ in inquisitive semantics is always *downward closed*. That is, if $\llbracket \varphi \rrbracket$ contains a proposition p then it must also contain any stronger proposition $q \subset p$. This is because, if p resolves the issue expressed by φ and does not contain any worlds that are ruled out by the information conveyed by φ , then the same must hold for any $q \subset p$. As a limit case, it is assumed that the inconsistent proposition, \emptyset , trivially resolves all issues, and is therefore included in the semantic value of every sentence. For any set of propositions \mathcal{P} we will write \mathcal{P}^\downarrow for the set of propositions q which are contained in some $p \in \mathcal{P}$:

$$(9) \quad \mathcal{P}^\downarrow := \{q \mid q \subseteq p \text{ for some } p \in \mathcal{P}\}$$

This allows for a compact notation of semantic values in inquisitive semantics. For instance, the semantic values of *Ann left* and *Did Ann leave?* can be written concisely as follows.

$$(10) \quad \llbracket \text{Ann left} \rrbracket = \{\{w \mid \text{Ann left in } w\}\}^\downarrow$$

$$(11) \quad \llbracket \text{Did Ann leave?} \rrbracket = \left\{ \begin{array}{l} \{w \mid \text{Ann left in } w\}, \\ \{w \mid \text{Ann didn't leave in } w\} \end{array} \right\}^\downarrow$$

Note that the *maximal elements* of $\llbracket \varphi \rrbracket$ are those propositions that contain precisely enough information to resolve the issue expressed by φ ; other propositions in $\llbracket \varphi \rrbracket$ resolve the issue expressed by φ as well, but contain additional information which may be irrelevant. For instance, the propositions $\{w \mid \text{Ann left in } w\}$ and $\{w \mid \text{Ann and Betty left in } w\}$ both resolve the issue expressed by *Did Ann leave?*. However, while the former contains precisely enough information to do so, the latter contains the additional information that Betty left. The maximal elements in $\llbracket \varphi \rrbracket$ are referred to in the inquisitive semantics literature as the *alternatives* that φ introduces. Declarative statements always introduce a single alternative. Questions on the other hand introduce multiple alternatives, and these alternatives correspond to the most basic ways of resolving the question. The use of the term ‘alternatives’ is

thus similar here to its use in Hamblin/alternative semantics (Hamblin 1973, Kratzer & Shimoyama 2002, Alonso-Ovalle 2006).⁴

Informative content and truth The set of all worlds that are compatible with the information that φ conveys is $\bigcup \llbracket \varphi \rrbracket$. This set of worlds is referred to as the *informative content* of φ and is denoted as $\text{info}(\varphi)$. For instance, the informative content of *Ann left* is the set of all worlds in which Ann left, and the informative content of *Did Ann leave?* is the set of all worlds in which Ann either left or didn't leave, that is, the set of all worlds whatsoever. In general, the informative content of a question is always trivial, in the sense that it always comprises all worlds that are compatible with what the question presupposes.

While in inquisitive semantics the semantic value of φ evidently does not correspond one-to-one to the truth-conditions of φ , it does *determine* these truth-conditions. Namely, φ is true in a world w just in case w is not ruled out by the information conveyed by φ . So φ is true in w if and only if $w \in \text{info}(\varphi)$, or formulated directly in terms of semantic content, if and only if $w \in \bigcup \llbracket \varphi \rrbracket$.

Informative and inquisitive sentences A sentence φ is called *informative* if and only if its informative content is non-trivial, i.e., $\text{info}(\varphi) \neq W$. Similarly, it is called *inquisitive* just in case the issue it expresses is non-trivial. This is the case if and only if in determining the semantic value of φ , condition (i) above is not trivially satisfied, i.e., if being an element of $\llbracket \varphi \rrbracket$ requires more than just consisting of worlds that are compatible with the information conveyed by φ . This holds if and only if $\text{info}(\varphi) \not\subseteq \llbracket \varphi \rrbracket$.

If φ is non-inquisitive, it introduces a unique alternative, namely $\text{info}(\varphi)$. Vice versa, if φ introduces multiple alternatives, it is inquisitive. For instance, *Ann left* is informative but non-inquisitive, while conversely, *Did Ann leave?* is inquisitive but non-informative.

These are all the notions from inquisitive semantics that are needed here: the semantic content of a sentence φ , $\llbracket \varphi \rrbracket$, is a downward closed set of propositions, always including \emptyset as a limit case; the alternatives that φ introduces are the maximal elements of $\llbracket \varphi \rrbracket$; $\text{info}(\varphi)$ amounts to $\bigcup \llbracket \varphi \rrbracket$; φ is true in w iff $w \in \text{info}(\varphi)$; φ is informative iff $\text{info}(\varphi) \neq W$; and φ is inquisitive iff $\text{info}(\varphi) \not\subseteq \llbracket \varphi \rrbracket$.

Using these notions, we now specify a semantic treatment of declarative and interrogative complements (§2.2) and our baseline entries for *believe* and *wonder* (§2.3-2.4).

⁴ Though see Ciardelli & Roelofsen (2017) for discussion of some subtle differences between these two notions of ‘alternatives’. These differences are crucial elsewhere but irrelevant for our purposes here.

2.2 Declarative and interrogative complements

We assume here that declarative and interrogative complements have the same semantic value as the corresponding matrix clauses. For instance, the declarative complement *that Ann left* has the same semantic value as the matrix declarative *Ann left*:

$$(12) \quad \llbracket \text{that Ann left} \rrbracket = \{ \{w \mid \text{Ann left in } w\} \}^\downarrow$$

Following [Ciardelli et al. \(2015\)](#) and much other work in inquisitive semantics, we assume here that a declarative complement or matrix clause ϕ (with falling intonation) is never inquisitive.⁵ That is, it always introduces a single alternative. This holds in particular for *disjunctive* declaratives, which will play a prominent role below. For example, the disjunctive complement *that Ann or Betty left* is taken to have the following semantic value:

$$(13) \quad \llbracket \text{that Ann or Betty left} \rrbracket = \{ \{w \mid \text{Ann or Betty left in } w\} \}^\downarrow$$

On the other hand, we assume that interrogative complements, just like matrix interrogatives, are never informative. This means that the propositions contained in the semantic value of an interrogative complement clause always completely cover the set of all possible worlds in which the presuppositions of the clause are satisfied.

The semantic values we assume for various kinds of interrogative complements are given in (14)-(17) below, and also depicted in Figure 1. The complement in (14) corresponds to our earlier matrix interrogative example *Did Ann leave?*.

$$(14) \quad \llbracket \text{whether Ann left} \rrbracket = \left\{ \begin{array}{l} \{w \mid \text{Ann left in } w\}, \\ \{w \mid \text{Ann didn't leave in } w\} \end{array} \right\}^\downarrow$$

⁵ There is also work in inquisitive semantics that does not make this assumption (e.g., [AnderBois 2012](#)). This requires a view under which uttering an inquisitive sentence does not necessarily involve issuing a request for information. See [Ciardelli et al. \(2012\)](#) for discussion.

EARLY ACCESS

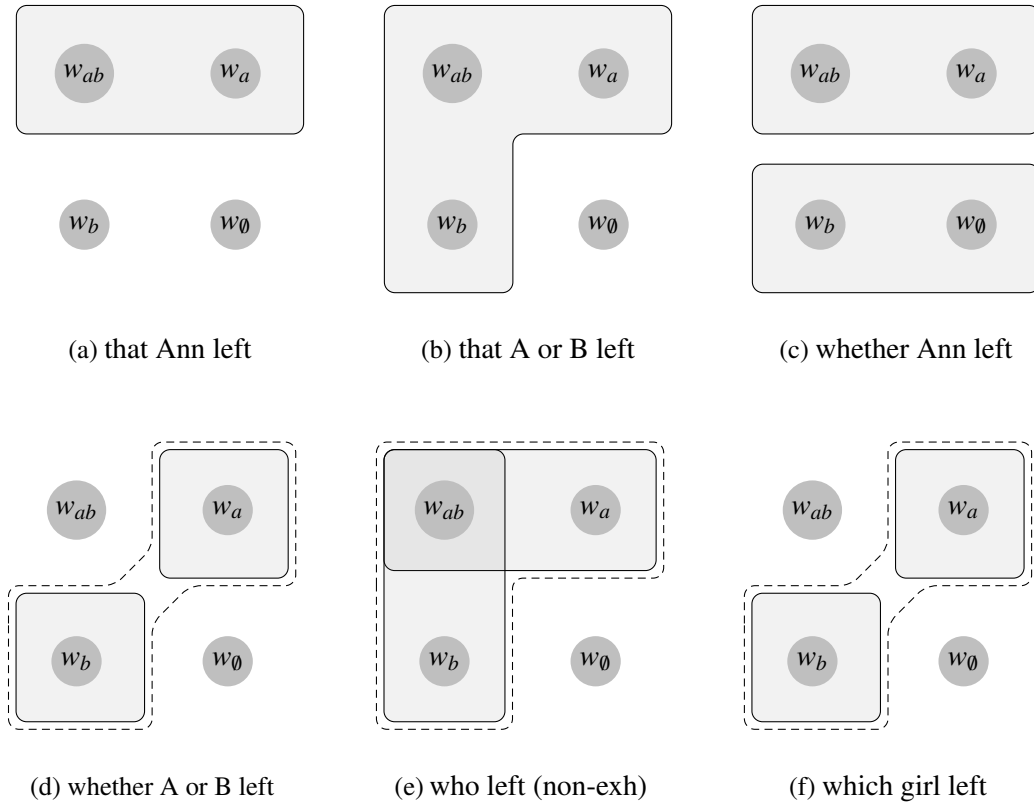


Figure 1 The semantic values of various kinds of complement clauses. In each diagram, w_{ab} is a world where both Ann and Betty left, w_a a world where only Ann left, w_b one in which only Betty left, and w_\emptyset one in which neither Ann nor Betty left. Non-trivial presuppositions are depicted with dashed lines. Trivial presuppositions are not depicted.

Next we consider the disjunctive complement *whether Ann or Betty left*. Such a disjunctive *whether*-complement has two possible interpretations. It may be interpreted as a polar question, expressing the issue whether the disjunction as a whole is true or not, or as an alternative question, expressing the issue which of the two disjuncts is true, presupposing that exactly one of them is. In matrix interrogatives, these two interpretations correspond to different intonation patterns (see, e.g., [Bartels 1999](#), [Biezma & Rawlins 2012](#), [Pruitt & Roelofsen 2013](#)), but in embedded contexts one and the same intonation pattern usually allows for both interpretations (presumably, the differences in intonation we observe in matrix cases are masked in embedded contexts by the prosody of the sentence that the embedded clause is part of). In (15)

below we specify the semantic value and the presupposition of *whether Ann or Betty left* under its alternative question interpretation. Its polar question interpretation is very similar to that of the complement considered in (14) above. We focus here on the alternative question interpretation because that is most relevant for our discussion below.

- (15) a. $\llbracket \text{whether Ann or Betty left} \rrbracket = \left\{ \begin{array}{l} \{w \mid \text{Ann left in } w \text{ and Betty didn't}\}, \\ \{w \mid \text{Betty left in } w \text{ and Ann didn't}\} \end{array} \right\}$
 b. $\text{presup}(\text{whether Ann or Betty left}) = \left\{ w \mid \begin{array}{l} \text{exactly one of Ann} \\ \text{and Betty left in } w \end{array} \right\}$

We now turn to *wh*-questions. First consider *who left*. Let us assume that the quantificational domain of the *wh*-phrase consists of just two individuals, Ann and Betty. Let us further assume, as is often done in the literature, that *wh*-questions come with an existential presupposition. Thus, *who left* presupposes that someone in the quantificational domain left, i.e., either Ann or Betty, or both.⁶ This yields the semantic analysis in (16):⁷

- (16) a. $\llbracket \text{who left} \rrbracket = \left\{ \begin{array}{l} \{w \mid \text{Ann left in } w\}, \\ \{w \mid \text{Betty left in } w\} \end{array} \right\}^\downarrow$
 b. $\text{presup}(\text{who left}) = \{w \mid \text{Ann or Betty left in } w\}$

Finally, consider the complement *which girl left*. We assume, again following much existing work, that this complement presupposes that exactly one girl left. We still take the quantificational domain of the *wh*-phrase to consist of Ann and Betty, so the presupposition is that exactly one of Ann and Betty left. This yields the semantic analysis in (17):

- (17) a. $\llbracket \text{which girl left} \rrbracket = \left\{ \begin{array}{l} \{w \mid \text{only Ann left in } w\}, \\ \{w \mid \text{only Betty left in } w\} \end{array} \right\}^\downarrow$
 b. $\text{presup}(\text{which girl left}) = \left\{ w \mid \begin{array}{l} \text{exactly one of Ann} \\ \text{and Betty left in } w \end{array} \right\}$

⁶ Nothing in our discussion below hinges on this assumption.

⁷ The semantic value assumed here captures the non-exhaustive (mention-some) reading of *who left*. The account can be refined to derive strongly and intermediate exhaustive readings as well (see Theiler et al. 2018). This refinement, however, doesn't affect any of the results presented here and is therefore omitted.

Having laid out a semantic analysis of declarative and interrogative complements which is representative for most analyses found in the literature,⁸ we are now ready to specify a baseline account of *believe* and *wonder*.

2.3 *Believe*

We assume, in line with the canonical treatment of *believe* in epistemic logic (Hintikka 1962) and much subsequent work, that the semantic value of a sentence of the form *x believes φ* is the set of all propositions *p* consisting of worlds *w* in which the doxastic state of *x*, DOX_x^w , coincides with a proposition in $\llbracket \varphi \rrbracket$.⁹ Moreover, following Karttunen (1974) and many others, we take *x believes φ* to presuppose that *x* believes the presuppositions of φ , i.e., that $\text{DOX}_x^w \subseteq \text{presup}(\varphi)$.¹⁰

- (18) a. $\llbracket x \text{ believes } \varphi \rrbracket = \{ \{w \mid \text{DOX}_x^w \in \llbracket \varphi \rrbracket \} \}^\downarrow$
 b. $\text{presup}(x \text{ believes } \varphi) = \{w \mid \text{DOX}_x^w \subseteq \text{presup}(\varphi)\}$

Consider the following example:

- (19) John believes that Ann left.

The semantic value of (19) is predicted to be the set of all propositions *p* consisting of worlds *w* in which John's doxastic state, DOX_j^w , is an element of $\llbracket \text{that Ann left} \rrbracket$. The latter means that DOX_j^w must only contain worlds in which Ann left. This implies that (19) is true in a world *w* if and only if all worlds in DOX_j^w are ones in which Ann left. These are indeed the desired truth conditions for the sentence. As for presuppositions, since the presupposition of the complement in (19) is trivial, i.e., $\text{presup}(\text{that Ann left}) = W$, the presupposition of the sentence as a whole is trivial as well.

Now consider a case with a presuppositional complement:

- (20) John believes that the king of France left.

⁸ Spelling out how the semantic values of these complements could be derived compositionally in inquisitive semantics would take us too far afield here; see Ciardelli et al. (2015), Roelofsen (2015), Champollion et al. (2015).

⁹ As usual, we take DOX_x^w to be the set of possible worlds that are compatible with *x*'s beliefs in *w*.

¹⁰ This canonical analysis of *believe* needs to be refined if we want to capture the fact that the verb is neg-raising and does not license interrogative complements. Such refinements can be found in Gajewski (2007), Romoli (2013), and Križ (2015), among others (for neg-raising), and in Theiler et al. (2017a,b), Mayr (2017), and Cohen (2017) (for the incompatibility with interrogative complements). As far as we can see, these refinements are orthogonal to the issues that concern us here and the arguments we will make.

The complement presupposes that there is a unique king of France. Thus, (20) as a whole presupposes that John believes there to be a unique king of France. Further, the semantic value of (20) is the set of all propositions p consisting of worlds w such that DOX_j^w consists exclusively of worlds in which the king of France left. This implies that (20) is true in a world w if and only if all worlds in DOX_j^w are ones in which the king of France left—again, these are the desired truth conditions, which form the core of all existing accounts of *believe*.

2.4 Wonder

To model what it means for an individual to *wonder* about something, we do not only need a formal representation of her doxastic state, but also a representation of the issues that she entertains, i.e., her *inquisitive state*. Following Ciardelli & Roelofsens (2015), we formally model an individual x 's inquisitive state in a world w , INQ_x^w , as a downward closed set of propositions which together cover her doxastic state, i.e., $\bigcup \text{INQ}_x^w = \text{DOX}_x^w$. The propositions in INQ_x^w are those that contain enough information to resolve the issues that x entertains. They correspond to extensions of x 's current doxastic state in which all her questions are settled one way or another.

Intuitively, x wonders about a question, e.g., about *who left*, just in case (i) x isn't certain yet who left, and (ii) she wants to find out who did. This is the case exactly if (i) x 's current doxastic state does not resolve the question yet; and (ii) every proposition/doxastic state in x 's inquisitive state is one that does resolve the question. Thus, Ciardelli & Roelofsens (2015) propose the analysis in (21a).¹¹ We add here that a sentence of the form x wonders ϕ presupposes that x believes the presuppositions of ϕ to be true, as captured by (21b).

- (21) a. $\llbracket x \text{ wonders } \phi \rrbracket = \{ \{w \mid \underbrace{\text{DOX}_x^w \not\subseteq \llbracket \phi \rrbracket}_{x \text{ isn't certain yet. } \dots} \wedge \underbrace{\text{INQ}_x^w \subseteq \llbracket \phi \rrbracket}_{\text{but wants to find out}} \} \}^\downarrow$
- b. $\text{presup}(x \text{ wonders } \phi) = \{w \mid \text{DOX}_x^w \subseteq \text{presup}(\phi)\}$

To illustrate the predictions that this analysis makes, first consider the following example, where the complement does not carry any presupposition.

- (22) John wonders whether Ann left.

¹¹ One attractive feature of this analysis is that it straightforwardly accounts for the fact that *wonder* does not take declarative complements. Namely, the analysis predicts that combining the verb with such a complement always yields a contradiction (Ciardelli & Roelofsens 2015, 2018). See also Uegaki (2015) for a very similar result, and Theiler et al. (2017a) for comparison.

The semantic value of the complement, $\llbracket \textit{whether Ann left} \rrbracket$, is the set of all propositions p such that either (i) all worlds in p are ones where Ann left, or (ii) all worlds in p are ones where Ann didn't leave, as was depicted in Figure 1(c). Therefore, the semantic value of (22) as a whole is the set of all propositions consisting of worlds w such that $\text{DOX}_j^w \not\subseteq \llbracket \textit{whether Ann left} \rrbracket$ and $\text{INQ}_j^w \subseteq \llbracket \textit{whether Ann left} \rrbracket$. The first requirement is satisfied just in case DOX_j^w contains at least one world in which Ann left and at least one world in which she didn't leave, i.e., in case John doesn't know yet whether Ann left. The second requirement is satisfied just in case every extension of John's current doxastic state in which the issues that he entertains are resolved is one in which he has come to know whether Ann left. This seems to be precisely what is expressed by (22).

Now consider an example with a presuppositional complement:

(23) John wonders which girl left.

Suppose, as before, that the quantificational domain of the *wh*-phrase consists of Ann and Betty. The predictions, then, are as follows. First, it is predicted that (23) presupposes that John believes that exactly one of Ann and Betty left. Further, it is predicted that the semantic value of (23) is the set of all propositions consisting of worlds w such that $\text{DOX}_j^w \not\subseteq \llbracket \textit{which girl left} \rrbracket$ and $\text{INQ}_j^w \subseteq \llbracket \textit{which girl left} \rrbracket$. The first requirement is satisfied just in case John doesn't know yet which of the two girls left. The second requirement is satisfied just in case every extension of John's current doxastic state in which the issues that he entertains are resolved is one in which he has come to know which girl left.

2.5 Distributive ignorance is not captured

The baseline entries for *believe* and *wonder* given above do not predict DIIs. To see this consider the following scenario.

(24) **Scenario** A crime has been committed, and there are three suspects: Ann, Bill and Carol. Detective Jones is on the case, and has already ruled out that Carol did it. However, he has not determined yet whether Ann or Bill did it.

In this scenario, consider sentences (1) and (2) from the introduction, repeated in (25) and (26) below:

(25) The detective wonders whether Ann, Bill, or Carol did it.

(26) The detective believes that Ann, Bill, or Carol did it.

These sentences seem false in the given scenario, because Jones has already ruled out the possibility that Carol did it (the experiments presented in Sections 4 and 5

show that this is indeed the majority judgment). However, our baseline entries for *believe* and *wonder* predict the sentences to be true.

To see this, let us first explicate what Jones' doxastic state and inquisitive state are in the described scenario. For simplicity, let us assume that our logical space consists of just four possible worlds, w_A , w_B , w_C , and w_E . In w_A , Ann committed the crime, in w_B Bill did it, in w_C Carol did it, and in w_E the crime was not committed by any of the current suspects but rather by someone else.

Jones' information state, DOX_j^w , consists of those worlds in which either Ann or Bill did it, i.e., $\text{DOX}_j^w = \{w_A, w_B\}$. This is depicted in Figure 2(a). On the other hand, Jones' inquisitive state, INQ_j^w , consists of all extensions of his current information state in which the issue that he entertains is resolved. These are the states $\{w_A\}$, $\{w_B\}$, and \emptyset (recall that it is assumed that the inconsistent information state, \emptyset , trivially resolves all issues). INQ_j^w is depicted in Figure 2(b).

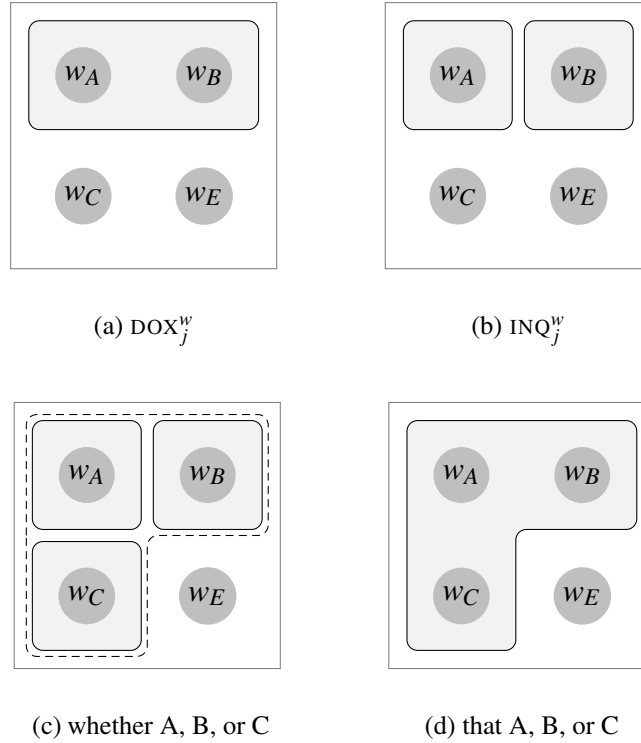


Figure 2 Jones' information state and inquisitive state in the scenario in (24), and the meaning of the complements in (25) and (26). Recall that inquisitive states and complement meanings are downward closed; only their maximal elements are depicted.

Let us now turn to the complements in (25) and (26). The interrogative complement in (25) expresses an issue whose resolution requires establishing which of Ann, Bill, and Carol did it, presupposing that one of them did. Thus, as depicted in Figure 2(c), its semantic value contains three alternatives, each corresponding to one of the three disjuncts. On the other hand, the declarative complement in (26) conveys the information that one of Ann, Bill, and Carol did it, and does not express an issue requesting any further information. Thus, as depicted in Figure 2(d), its semantic value contains a single alternative consisting of all worlds in which Ann, Bill, or Carol did it.

The entry for *believe* predicts that (26) is true if and only if the following requirement is met:

$$(27) \quad \text{DOX}_j^w \in \llbracket \text{that A, B, or C} \rrbracket$$

By inspecting Figures 2(a) and 2(d), it can be seen that this requirement is indeed met. So (26) is incorrectly predicted to be true.

On the other hand, the entry for *wonder* predicts that (25) is true if and only if the following requirements are met:

$$(28) \quad \begin{array}{ll} \text{a.} & \text{DOX}_j^w \subseteq \{w_A, w_B, w_C\} \\ \text{b.} & \text{DOX}_j^w \notin \llbracket \text{whether A, B, or C} \rrbracket \\ \text{c.} & \text{INQ}_j^w \subseteq \llbracket \text{whether A, B, or C} \rrbracket \end{array} \quad (\text{presupposition})$$

By inspecting Figures 2(a)-2(c), it can be seen that these requirements are met as well. So (25) is also incorrectly predicted to be true.

3 Toward a refined account: two parameters

There are various ways to refine the assumed baseline account of *believe* and *wonder* so as to derive DIIs. Of course, the choice between these various options cannot be made merely on the basis of the two simple example sentences we have considered so far. Rather, to provide a richer testbed, we should investigate variants of these basic cases as well. Before turning to this empirical investigation, however, we should determine which kinds of constructions to look at exactly. This requires a preliminary exploration of the various theoretical options. Thus, our aim in this section will not be to spell out any particular theory in full detail, but rather to discuss the main parameters that set the various approaches apart.

There are, as far as we can see, two main parameters, which we will label *structure-sensitivity* and *locality*, respectively. We will discuss each in turn.

3.1 Structure-sensitivity

Compare the following two pairs of sentences.

- (29) a. The detective wonders whether Ann, Bill, or Carol did it.
 b. The detective wonders which of the suspects did it.
- (30) a. The detective believes that Ann, Bill, or Carol did it.
 b. The detective believes that one of the suspects did it.

Note that (29a) is a repetition of our initial example (1), and that (29b) is a variant of it in which the verb takes a *wh*-question as its complement rather than an alternative question. Similarly, (30a) is a repetition of our earlier example (2), and (30b) is a variant of it in which the complement involves an existential quantifier rather than a disjunction.

In the scenario under consideration, where the suspects are Ann, Bill, and Carol, the two interrogative complements in (29a) and (29b) are semantically equivalent, even though they are structurally different, and the same goes for (30a) and (30b). The various possible refinements of our baseline account can be divided into ones that are sensitive to such structural differences and ones that are not. Note that the baseline account itself is not sensitive to the structure of the clauses that *wonder* and *believe* take as their complement. According to the given entries, the verbs only operate on the semantic content of their complement. Thus, in the absence of any further assumptions, it is predicted that (29a) and (29b) are equivalent, and similarly for (30a) and (30b). Importantly, this does not only hold for the baseline account itself, but also for variants of it that remain insensitive to the structure of the clauses that the verbs take as their complement. This includes any account which derives DIIs as pragmatic implicatures and which assumes that the formal alternatives that play a role in the computation of such implicatures are fully dictated by the semantic content of the sentence under consideration and contextual factors such as the question under discussion.¹²

An example of a structure-sensitive account would be one that derives DIIs as pragmatic implicatures but does *not* assume that the relevant formal alternatives are fully determined by semantic content and contextual factors such as the question under discussion, but also depend on the structure of the sentence involved. A general argument for such a structure-sensitive approach to computing implicatures has been

¹² A terminological note: here and below, we use the term ‘formal alternatives’ for the expressions that are taken into consideration when computing the implicatures of a given expression φ . These formal alternatives are not to be confused with the maximal elements of $\llbracket \varphi \rrbracket$, which are referred to in inquisitive semantics and Hamblin semantics as the alternatives that φ introduces. Note that the former are syntactic objects while the latter are semantic objects. When referring to the former kind of alternatives we will always speak of *formal* alternatives.

made by [Katzir \(2007\)](#). In particular, he proposes that the formal alternatives of a sentence φ that are taken into account when computing implicatures are only those sentences that can be obtained from φ either (i) by deleting elements in φ , or (ii) by substituting elements in φ with other elements from an appropriately defined source. Such an account would predict a difference in interpretation between (30a) and (30b), as follows. In the case of (30a), pragmatic reasoning would involve the formal alternatives in (31) below, which can all be obtained from (30a) by deleting parts of it:

- (31) a. The detective believes that Ann or Bill did it.
- b. The detective believes that Ann or Carol did it.
- c. The detective believes that Bill or Carol did it.
- d. The detective believes that Ann did it.
- e. The detective believes that Bill did it.
- f. The detective believes that Carol did it.

These formal alternatives all entail (30a) itself. Thus, routine pragmatic reasoning leads to the conclusion that, if a speaker utters (30a), he does not have enough information to affirm any of the formal alternatives in (31), and if we further assume that the speaker is knowledgeable about what the detective believes, we end up deriving that none of the formal alternatives in (31) are true. This, in turn, implies that the detective must still consider all of Ann, Bill, and Carol possible culprits. Thus, the DII is accounted for.

In the case of (30b), on the other hand, this derivation does not get off the ground, because the formal alternatives in (31) cannot be obtained from (30b) by deletion and are therefore not taken into account when computing implicatures. Thus, it would be predicted that (30b) does not imply distributive ignorance.

[Roelofsen & Uegaki \(2016\)](#) suggest, based on introspective judgments collected from a small group of informants, that these predictions are correct: (30a) gives rise to a DII, but (30b) does not. However, they also point out that it may in principle be possible to account for this contrast under the assumption that the mechanism responsible for DIIs is only sensitive to the semantic content of the complement clause, and not (or at least not directly) to its syntactic structure. Namely, the fact that a DII is absent in (30b) may be due to implicit *domain restriction*. For instance, (30b) can perhaps be interpreted as follows:

- (32) The detective believes that one of the [most likely] suspects did it.

[Roelofsen & Uegaki \(2016\)](#) discuss various challenges for such an account. Here, we will focus on a particular empirical prediction, which, again in the absence of further assumptions, sets the domain restriction account apart from structure-sensitive ones.

This prediction concerns sentences like (33) and (34), which only differ from (29b) and (30b), respectively, in that they involve the numeral *three*.

(33) The detective wonders which of the three suspects did it.

(34) The detective believes that one of the three suspects did it.

Geurts & van Tiel (2016) provide experimental evidence that implicit domain restriction is very unlikely in partitive structures with a numeral, such as the *wh*-phrase in (33) and the existential quantifier in (34). Thus, an account which assumes that DIIs are not structure-sensitive and predicts a contrast between (29a) and (29b) based on the possibility of domain restriction in (29b), predicts that (33), where domain restriction is blocked by the numeral, will pattern with (29a) rather than with (29b). That is, on such an account (33) is expected to give rise to a DII.

On the other hand, a structure-sensitive account of DIIs such as the one sketched above, which predicts a contrast between (29a) and (29b) based on structural differences, leads us to expect, in the absence of further assumptions, that (33) will pattern with (29b) rather than with (29a). That is, on such an account (33) is in principle predicted *not* to give rise to a DII.

Roelofsen & Uegaki (2016) tentatively suggest, again based on introspective judgments of a small group of informants, that (33) does give rise to DIIs. However, the judgments were rather mixed in this case.

In our Experiment 1, to be discussed below, we tested the structure-sensitivity of DIIs by comparing sentences involving four types of complements: disjunctive complements (as in (29a)), quantificational complements whose domain is specified by a noun phrase without a numeral (as in (29b)), quantificational complements whose domain is specified by a noun phrase with a numeral (as in (33)), as well as quantificational complements whose domain is listed explicitly by means of a conjunction (as in *The detective wonders which of Ann, Bill and Carol did it*). If DIIs are structure-sensitive, the latter are expected to pattern with disjunctive complements.

3.2 Locality

We now turn to the second important parameter on which the various possible refinements of the assumed baseline account would differ. To see what this amounts to, consider the following scenario:¹³

(35) There is a crime with three suspects, Ann, Bill, and Carol. There are three detectives investigating the case.

¹³ We thank Benjamin Spector for drawing our attention to this type of scenarios.

- Detective 1 still considers Bill and Carol possible culprits, but not Ann.
- Detective 2 still considers Ann and Carol possible culprits, but not Bill.
- Detective 3 still considers Ann and Bill possible culprits, but not Carol.

In this scenario, consider the following two sentences:

(36) Every detective is wondering whether Ann, Bill, or Carol did it.

(37) Every detective believes that Ann, Bill, or Carol did it.

Note that these are variants of our initial examples (1) and (2) in which the subject is a universal quantifier rather than a referential expression. Consider the structure-sensitive account of DIIs as pragmatic implicatures outlined above. When we apply this account to (37), the following implicatures are derived:

- (38) It is not the case that every detective believes. . .
- a. . . that Ann or Bill did it.
 - b. . . that Ann or Carol did it.
 - c. . . that Bill or Carol did it.
 - d. . . that Ann did it.
 - e. . . that Bill did it.
 - f. . . that Carol did it.

These implicatures are all true in the given scenario, and the same holds for the literal meaning of the sentence according to our baseline entry. Thus, the sentence is predicted to be true, even though in the given scenario all the detectives have already ruled out one of the suspects, so none of them is distributively ignorant.

Besides pragmatic theories which predict that the literal interpretation of an utterance may be strengthened through reasoning about formal alternatives of the uttered sentence as a whole, there are also theories which predict that the meaning of any *part* of a given sentence may be strengthened through comparison with formal alternatives, and that this strengthened meaning may then serve as input to operators that apply to this part of the sentence, all within the process of composing the semantic content of the sentence, before pragmatic reasoning enters the stage (see, e.g., Chierchia et al. 2012). This process of strengthening through comparison with formal alternatives is referred to as *exhaustification*.

Thus, while pragmatic reasoning can only have a ‘global’ effect on the interpretation of a sentence as a whole, exhaustification as conceived by Chierchia et al. (2012) and others can also have a ‘local’ effect, on parts of a sentence. For instance, if exhaustification is applied to the verb phrase in (37), before the universal quantifier in subject position is composed with it, we derive the following implications:

- (39) Every detective is such that it is not the case that she believes. . .
- a. . . that Ann or Bill did it.
 - b. . . that Ann or Carol did it.
 - c. . . that Bill or Carol did it.
 - d. . . that Ann did it.
 - e. . . that Bill did it.
 - f. . . that Carol did it.

Notice that the first three implications are false in the given scenario. Thus, on an account which assumes that the meaning of the verb phrase in (37) is *obligatorily* strengthened through exhaustification, the sentence is predicted to be false, unlike on the pragmatic account considered above. More specifically, under the assumption that local exhaustification is obligatory, (37) is only true in situations in which every detective is distributively ignorant. On the other hand, if exhaustification of the verb phrase is only considered *optional*, then (37) is predicted to have two readings: one under which it is true in the given scenario (without local exhaustification), and one under which it is false (with local exhaustification).

Which of these two readings is preferred could then be taken to depend on pragmatic factors (Chierchia et al. 2012, Potts et al. 2016). In particular, Chierchia et al. (2012) suggest that some version of Dalrymple et al.'s (1998) Strongest Meaning Hypothesis plays an important role in determining such preferences. According to this suggestion, other things being equal, a reading involving local exhaustification is preferred if it is stronger than the reading obtained without local exhaustification, and dispreferred if it is weaker than that reading. Thus, in the case of (37) for instance, the reading with local exhaustification would be preferred, since it is stronger than the one without. However, if the quantifier in subject position were downward entailing rather than upward entailing (e.g., *no detective* rather than *every detective*) the reading without local exhaustification would in principle be preferred. Finally, if the quantifier is non-monotonic (e.g., *exactly two detectives*), local exhaustification makes the reading neither stronger nor weaker. This means that the Strongest Meaning Hypothesis typically does not constrain local exhaustification under non-monotonic quantifiers.

Let us now take a step back. Given the considerations above, the various possible refinements of our baseline account that derive the DIIs of (1) and (2) through comparison with formal alternatives can be of three types: they could assume that such comparison (i) only happens globally, (ii) optionally happens locally as well, or (iii) obligatorily happens locally. The latter type of theory can be implemented, for instance, by incorporating an exhaustification operator in the lexical semantics of the relevant verbs, as is done for *wonder* (though not *believe*) in Roelofsen & Uegaki (2016).

	Structure-sensitive	Not structure-sensitive
Only global strengthening	Approach 1	Approach 4
Optional local strengthening	Approach 2	Approach 5
Obligatory local strengthening	Approach 3	Approach 6

Table 1 Overview of possible theoretical approaches.

Combining the two parameters we have considered, structure-sensitivity and locality, we can distinguish six general theoretical approaches, as indicated in Table 1. In the following two sections, we will report the results of two experiments that were aimed to determine which of these general approaches is most adequate. This is done by considering variants of our basic examples in which the structure of the complement clause is different (Experiment 1), and ones involving different kinds of quantificational subjects (Experiment 2). The former should allow us to tease apart structure-sensitive approaches from structure-insensitive ones. The latter should allow us to distinguish between approaches that assume only global strengthening, and ones that assume optional or obligatory local strengthening through comparison with formal alternatives.

We should note that we are not presupposing at this point that one and the same approach is most fitting to account for DIIs of both *wonder* and *believe*. In fact, in Roelofsen & Uegaki (2016: footnote 5) it was explicitly suggested that Approach 3 in Table 1 is most appropriate for *wonder*, while Approach 1 is most suitable for *believe*. Anticipating what is to come, the results of our experiments refute this hypothesis and suggest instead that Approach 2 is most adequate for both *wonder* and *believe*.

4 First experiment: complements

4.1 Goal

The goal of our first experiment was to compare the DIIs with *wonder* and *believe*, and to investigate the potential contrast illustrated in (29) and (30). In particular, we wanted to test the claim made by Roelofsen & Uegaki (2016) that there is a contrast between, on the one hand, disjunctive complements (‘whether/that A, B, or C’) and complements whose domain is explicitly listed by means of a conjunction over individuals (‘which/one of A, B, and C’) and, on the other hand, complements whose domain is not explicitly listed (‘which/one of the suspects’). Another case of interest discussed in §3.1 is that of numerals. Since it is known that numerals in the domain phrase (‘which/one of the *three* suspects’) block implicit domain restriction

(Geurts & van Tiel 2016), we wanted to see whether the presence of numerals would affect the judgments concerning DIIs.

4.2 Design

Participants were recruited on Amazon’s Mechanical Turk and took a survey directly on the platform. The survey consisted of the context in (40) and three sentences they had to judge. Each sentence was followed by the question “In this context, would you say that this sentence is true or false?” and a 7-point scale, the extreme points of which were labeled ‘Clearly false’ and ‘Clearly true’. These three sentences were followed by two demographic questions, asking for participants’ age and native language (a text field that participants could fill in freely). The whole survey was presented on a single page.

- (40) **Context:** Sue has three children, Sophie, Bill, and Mary, who all live on their own. Sue is impatiently waiting for all of them to arrive at her place for Thanksgiving dinner. Someone rings the bell. Sue isn’t sure who it is, but she knows that it can’t be Bill, because he just texted her that he would be late.

We recruited a relatively high number of participants (384) but kept the number of items per participant low, for the following reasons. First, presenting more items to each participant would have involved introducing multiple background stories, which would have made the survey more tedious and possibly confusing. Moreover, by keeping the number of items per participant low we obtained a good view of individual differences, and each participant was more naive to the goal of the experiment. Finally, having an order of magnitude more participants than in usual psycholinguistic experiments also has the advantage of facilitating the convergence of complex mixed-effects models (see Phillips & George 2018 for a similar method applied to the no-false-beliefs inference of embedded mention-some questions).

Eight different surveys were designed, corresponding to the eight possible target sentences obtained by combinations of the following two factors: embedding VERB (‘wonder’ or ‘believe’) and COMPLEMENT type (disjunction, conjunction, NP, NumP). Each survey version had 3 sentences (one for each CONDITION: Target, True control, and False control), which were presented in random order. Target sentences were literally true, but their DII was false in the given context. This way, we could test whether participants had derived the inference by measuring how unacceptable the target was for them. The target sentences for each combination are listed in Table 2. The True sentences were derived by dropping Bill from the disjunction and conjunction, and replacing ‘children’ or ‘three children’ with ‘daughters’ (this way, the DII was satisfied). The False sentence was the same across all four complement

types, and involved ‘whether Bill arrived’ and ‘that Bill arrived’ for *wonder* and *believe* respectively.

COMPLEMENT	VERB	Target sentence
Disjunction	<i>believe</i>	Sue believes that Sophie, Bill, or Mary arrived.
	<i>wonder</i>	Sue wonders whether Sophie, Bill, or Mary arrived.
Conjunction	<i>believe</i>	Sue believes that one of Sophie, Bill, and Mary arrived.
	<i>wonder</i>	Sue wonders which of Sophie, Bill, and Mary arrived.
NP	<i>believe</i>	Sue believes that one of her children arrived.
	<i>wonder</i>	Sue wonders which of her children arrived.
NumP	<i>believe</i>	Sue believes that one of her three children arrived.
	<i>wonder</i>	Sue wonders which of her three children arrived.

Table 2 Target sentences by COMPLEMENT and VERB.

4.3 Participants

For each survey version, we recruited eight participants for each of the six possible orders for the Target, True and False sentences (hence a total of 384 HITs were posted on Mechanical Turk, paid 27¢ each). Despite requesting that participants take the survey only once (and enforcing this with the UniqueTurker script), a few participants took multiple surveys. Three participants who took the survey 5 or more times were not paid for their retakes and their HITs were offered to new participants. All other participants were paid for all their retakes, but the corresponding data were discarded from the analyses. The data from 10 participants who reported native languages other than English were discarded as well.

In all, after removing all retakes and non-native speakers, we had data from 326 unique native English speakers (age range: 19–70).

4.4 Results

Anonymized data and analysis scripts for all experiments reported in this paper can be found at <https://semanticsarchive.net/Archive/TdhYjRiM/Cremers-Roelofsen-Uegaki.html>.

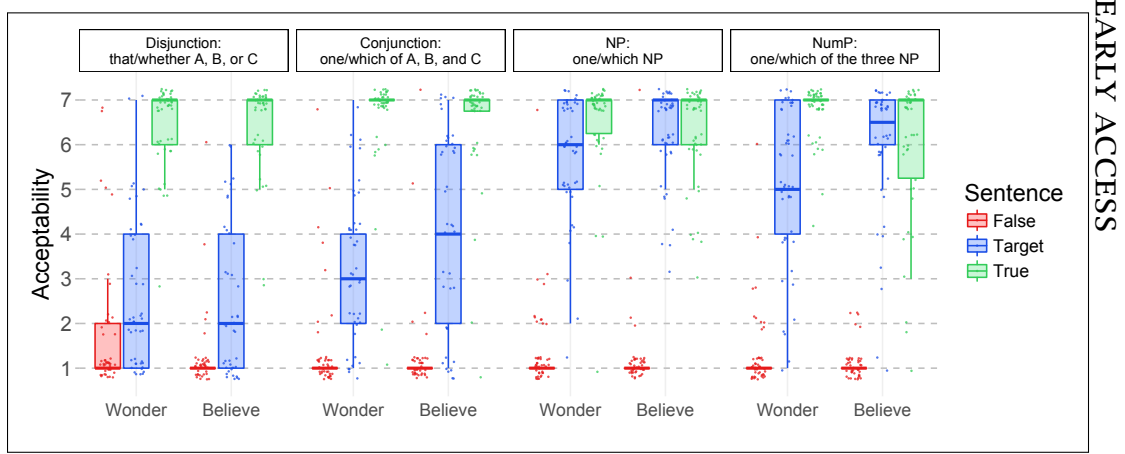


Figure 3 Experiment 1: Acceptability for each Verb, Complement, and Sentence (boxes indicate median and quartiles, each dot represents an individual answer).

The results are presented in Figure 3. Notice that DIIs are most visible with Disjunction and Conjunction COMPLEMENTS (as indicated by maximal difference between Targets and True controls), and seem to be completely absent with NP and NumP under *believe*.

We ran a mixed-effects ordinal logistic regression on the responses to the True and Target sentences with VERB (sum-coded), COMPLEMENT (baseline: Disjunctive), CONDITION (baseline: True) and all their interactions as fixed effects, and a random Subject intercept (in R, using package *ordinal*, R Core Team 2014, Christensen 2015). The full results are presented in Table 3. The highly significant effect of CONDITION confirms that Disjunction — the baseline for COMPLEMENT — gives rise to DIIs, and the absence of a $\text{CONDITION} \times \text{VERB}$ interaction indicated that this effect is similar for *wonder* and *believe*. That CONDITION interacts with COMPLEMENT on its NP and NumP levels shows that the DIIs have a very reduced effect with these two complements, although the triple interactions $\text{CONDITION} \times \text{VERB} \times \text{COMPLEMENT}$ indicate that unlike *believe*, *wonder* still gives rise to DIIs with these complements.

Model comparisons showed that there was no significant difference between the NP and NumP COMPLEMENTS ($\chi^2(4) = 4.2, p = .37$).

	β	z -value	p	
CONDITION	-4.49	-11.8	< 0.001	***
VERB	-0.01	-0.02	0.985	
[COMP:Conjunction]	0.55	1.5	0.131	
[COMP:NP]	-0.05	-0.1	0.888	
[COMP:NumP]	0.02	0.07	0.947	
CONDITION \times VERB	0.22	0.4	0.718	
CONDITION \times [COMP:Conjunction]	0.68	1.5	0.140	
CONDITION \times [COMP:NP]	3.65	7.8	< 0.001	***
CONDITION \times [COMP:NumP]	3.15	6.7	< 0.001	***
VERB \times [COMP:Conjunction]	0.37	0.5	0.610	
VERB \times [COMP:NP]	0.83	1.3	0.210	
VERB \times [COMP:NumP]	1.57	2.3	0.022	*
CONDITION \times VERB \times [COMP:Conjunction]	-1.51	-1.7	0.098	
CONDITION \times VERB \times [COMP:NP]	-1.75	-2.0	0.041	*
CONDITION \times VERB \times [COMP:NumP]	-2.90	-3.3	< 0.001	***

Table 3 Experiment 1: Results of the ordinal mixed-effects model on True and Target sentences. Factors of theoretical interest are highlighted for readability. The standard deviation of the random intercepts was .49. Significance levels: .05 (*), .01 (**), .001 (***).

4.5 Replications with different contexts

The experiment presented above involved only one context and the results we obtained may partly depend on certain features of this context. To address this issue, we ran two follow-up experiments with different contexts. We also made some small changes to the design. First, the VERB factor was made within-subject, so each participant saw two targets, two true controls and two false controls (one for each verb). Second, we added two fillers, one clearly true and one clearly false, bringing the total number of items a participant saw to eight (from three in the original experiment).

Details about these two follow-up experiments can be found in Appendix A. The key result was replicated (DIIs with disjunctive and conjunctive complements, but not with NP complements). However, in both experiments we observed a difference between NP and NumP complements. In one case, the NumP complements gave rise to DIIs as strong as disjunctive and conjunctive complements, and in the other case, it was somewhere in between. Furthermore, the effect of VERB on NumP targets was not replicated in one of the two contexts.

4.6 Discussion

First, we observed clear differences between the various complements. More specifically, the two complements which mentioned each alternative (Disjunction and Conjunction) showed much stronger DIIs than the noun phrase complement. This is in line with the hypothesis that structurally determined formal alternatives play an important role (Katzir 2007, Fox & Katzir 2011), and that mentioned alternatives are strongly activated, as argued in various empirical domains (see e.g., Onea & Steinbach 2012, Starr 2014 for the role of mentioned alternatives in conditionals; Roelofsen et al. 2016 in embedded questions; Coppock & Brochhagen 2013 in modified numerals, and Csipak & Zobel 2014, Rojas-Esponda 2014 in discourse particles.)

Second, we only found small differences between *wonder* and *believe*, and these differences were not replicated in the follow-up experiments. Crucially, there was no difference at all between the two verbs in the Disjunction condition, where the strongest DIIs were observed.¹⁴

Finally, the addition of a numeral had very different effects in the different contexts we tested. In the original experiment, it had no effect with *believe* and if anything, only slightly strengthened DIIs with *wonder*. In the follow-ups it had a stronger effect, with both verbs. The finding that NumP did not pattern with Disjunction/Conjunction across contexts is important in view of the role of implicit domain restriction. Although the contrast between Disjunction/Conjunction and NP could in principle be explained in terms of the possibility of implicit domain restriction in the latter, this explanation is implausible in the case of NumP, given that implicit domain restriction is highly unlikely with a numeral (Geurts & van Tiel 2016).¹⁵ To sum up, the addition of a numeral can make DIIs as strong as with

¹⁴ Note that the three other complements do not offer as minimal a comparison as Disjunction, because they involve ‘which of *X*...’ with *wonder* but ‘that one of *X*...’ with *believe*. This could have affected the results in two ways. First, it opens the possibility of a specific reading with *believe* if ‘one’ somehow gets wide scope. When designing the experiment, we tried to block this reading as much as possible by making clear that Sue doesn’t know which of her daughters will arrive first, but the fact that true controls received a slightly lower rating with *believe* than with *wonder* could indicate that some participants still had the specific reading. Second, it may be that *which* activates domain alternatives in a way that *one* does not. This could lead to stronger DIIs with ‘wonders which of’ than with ‘believes that one of’.

¹⁵ There is one subtlety to note here: strictly speaking, Geurts & van Tiel (2016) show that numerals block domain restriction only in the case of *non-intersective* quantifiers (i.e. quantifiers *Q* such that the truth-conditions of “*Q A B*” are not entirely determined by $A \cap B$). It is not straightforward to say whether or not *wh*-phrases like *which of her children* are intersective in the relevant sense. In particular, interpreting the question that a *wh*-phrase is part of as strongly exhaustive or weakly exhaustive leads to different conclusions as to whether the *wh*-phrase is intersective or not.

disjunctions or completely absent; its effect depends on the context and sometimes on the embedding verb.

At this point, the results favor an approach that is structure-sensitive (§3.1), but they do not tell us anything yet about locality (§3.2) since all sentences had non-quantificational subjects. In the next experiment, we will focus on disjunctive complements, which offer both the most minimal comparison between *believe* and *wonder* and exhibit the strongest DII effects, and we will test how DIIs project from the scope of various quantified subjects.

5 Second experiment: quantified subjects

5.1 Goal

The goal of the second experiment was to investigate the projection of DIIs under quantified subjects. In principle, DIIs could always be computed globally, always locally, or sometimes locally and sometimes globally. In §5.2 we provide some empirical and theoretical background on local strengthening.

Note that the DIIs of *wonder* and *believe* could in principle give rise to different projection patterns in quantified cases. This would suggest that the DIIs of the two verbs are of a different nature and require separate treatment, as suggested in Roelofsen & Uegaki (2016).

5.2 Background on local strengthening

Chierchia (2004) argued that scalar implicatures — most notably the implicature from *some* to ‘not all’ — are sometimes computed locally (in the scope of quantifiers, negation, modals, etc.). He proposed that such local implicatures are due to a silent *exhaustification operator* which can be freely inserted in embedded positions. This proposal was met with strong criticism; see in particular Horn (2006) and Geurts (2009) for re-evaluations of Chierchia’s original motivating examples in a Neo-Gricean framework. New arguments for Chierchia’s proposal were offered in Chierchia et al. (2009, 2012).

Most importantly for us, this debate moved to experimental grounds with the publication of Geurts & Poussoulous (2009). In this study, the authors tested the availability of local exhaustification readings using a sentence-picture verification task. While they did not find any evidence for such readings, this sparked a flurry of experimental work using various methods, and there is now a rather broad consensus that these readings are in fact available (Clifton & Dube 2010, Chemla & Spector 2011, Benz & Gotzner 2014, Potts et al. 2016; see van Tiel et al. 2018 for a review and methodological discussion).

As a consequence, the debate has now moved from a dispute over the facts to a debate on how these facts should be accounted for. While Geurts & van Tiel (2013) maintain a pragmatic account complemented with a mechanism of truth-conditional narrowing which requires narrow focus, proponents of the grammatical theory maintain that local readings are derived by insertion of an embedded exhaustification operator. Finally, Bergen et al. (2016) propose an account based on lexical ambiguity for *some* and other lexical items (equivalent to the possibility of local exhaustification) and probabilistic Gricean style reasoning to resolve ambiguities (see Potts et al. 2016 for experimental evaluation of this model).

One thing that has become clear about the availability of local strengthening is that it greatly depends on the monotonicity of the embedding operator. Its availability in upward-entailing environments (e.g., under *every*) varies quite a bit from one experiment to the next. On the other hand, it tends to be low in downward-entailing environments (e.g., under *no*), and high under non-monotonic quantifiers such as *exactly n*.¹⁶

5.3 Design

In this experiment, we followed Roelofsen & Uegaki (2016) and the literature on local strengthening in testing the quantifiers *every*, *no* and *exactly n*. We did however stick to the experimental paradigm of Experiment 1 where sentences were judged against a context given by short vignettes, since the sentence-picture verification tasks that are often used in the literature on local strengthening are not well suited to test attitude reports.

We focused on disjunctive complements, which made DIIs easiest to detect, and tested three quantifiers as subject: *every* (upward-entailing), *no* (downward-entailing), and *exactly two* (non-monotonic). Six different surveys were designed, corresponding to the six combinations of VERB (*wonder* or *believe*) and subject QUANTIFIER (*every*, *no*, *exactly two*). As in Experiment 1, each survey involved a context, followed by a Target sentence, a True control sentence, and a False control sentence which were presented in random order.

We designed contexts which would make local and global DIIs come apart for each of the quantifiers. The context used in Experiment 1 would not have translated very well with quantified subjects, so we moved to contexts with detectives, which also happened to be closer to the examples of Roelofsen & Uegaki (2016). Context (41) below was used for *every* and *no*, while context (42) was used for

¹⁶ van Tiel et al. (2018) identify a number of low-level factors that may have affected the results of the various experiments that have been carried out and may partly explain the differences found among these results (e.g., the length of the experiment, the presence of the scalar alternatives in filler items, and the distribution of individuals satisfying the local implicature in the domain of the quantifier).

exactly two. Since the contexts became more complex, we made a few changes intended to help participants process the information given to them. First, we used colors to identify each suspect, so it was easier to retrieve the relevant piece of information in the context if a participant forgot what each detective knows about each person (remember that this was a one-page survey so the background story remained on screen). Second, we did not give all suspects a name (which would be difficult for participants to keep track of) but rather identified them in terms of their occupations (the butler, the gardener, et cetera).

- (41) ***Every/no Context:*** The rich lord Edgware has been murdered, and three detectives are investigating the case independently. The suspects are four people working for the lord: his butler, his gardener, his maid, and his cook. Every detective quickly established that the gardener is innocent. The first detective further established that the maid cannot be the culprit, while the two other detectives gathered evidence showing that it cannot be the cook. No detective is aware of the others' discoveries, so the first detective still considers the cook as a suspect, and the two others still consider the maid as a suspect.
- (42) ***Exactly two Context:*** The rich lord Edgware has been murdered, and three detectives are investigating the case independently. The suspects are four people working for the lord: his butler, his gardener, his maid, and his cook. Every detective quickly established that the gardener is innocent. The first detective further established that the maid cannot be the culprit. No detective is aware of the others' discoveries, so two of them still consider the maid as a suspect.

The crucial information in the two contexts is summarized in Table 4. Note that we did not use the context given in (35) because we wanted to make sure that the different detectives' beliefs would not contradict each other (there is at least one suspect that could be the murderer according to all of them), and that there was one suspect that had been cleared by all detectives (this was useful for control items).

The False, Target, and True test sentences for *wonder* are presented in (43a-45a), (43b-45b), and (43c-45c) respectively. For the targets, we indicate what the local and global DII amount to in each case.

- (43) *Every-wonder* sentences:
- a. Every detective wonders whether or not the cook committed the crime.
 - b. Every detective wonders whether the maid, the cook, or the butler committed the crime.

- (i) Global DII: Every detective still wonders which suspect committed the crime, and none of the three mentioned suspects has been cleared by all detectives.
 - (ii) Local DII: None of the three mentioned suspects has been cleared by any detective.
 - c. Every detective wonders whether or not the **butler** committed the crime.
- (44) *No-wonder* sentences:
- a. No detective wonders whether or not the **maid** committed the crime.
 - b. No detective wonders whether the **maid**, the **cook**, or the **butler** committed the crime.
 - (i) Literal reading:¹⁷ Every detective figured out who committed the crime.
 - (ii) Local DII: Every detective either figured out who committed the crime, or still considers all the three mentioned suspects.
 - c. No detective wonders whether or not the **gardener** committed the crime.
- (45) *Exactly two-wonder* sentences:
- a. Exactly two detectives wonder whether or not the **gardener** committed the crime.
 - b. Exactly two detectives wonder whether the **maid**, the **cook**, or the **butler** committed the crime.

¹⁷ Note that in this case global strengthening is vacuous. The literal reading we predict is in fact disjunctive: every detective either knows already who committed the crime or doesn't want to know. Given the context, we can rule out the second disjunct (assuming every detective is in fact interested in solving the case).

		butler	gardener	maid	cook
(41)	D1	suspect	innocent	innocent	suspect
	D2	suspect	innocent	suspect	innocent
	D3	suspect	innocent	suspect	innocent
(42)	D1	suspect	innocent	innocent	suspect
	D2	suspect	innocent	suspect	suspect
	D3	suspect	innocent	suspect	suspect

Table 4 Summary of the situation in the two contexts. For each detective, it is indicated who they consider to be innocent and who they still consider to be a possible culprit.

- (i) Global DII: Exactly two detectives haven't figured out who committed the crime yet, and none of the three mentioned suspects has been cleared by both of these two detectives.
- (ii) Local DII: Exactly two detectives still consider all the three mentioned suspects, while the third detective has ruled out at least one of them.
- c. Exactly two detectives wonder whether or not the **maid** committed the crime.

The judgment for each target item in its respective context depended on whether a DII was derived locally, globally, or not derived at all. However, the exact mapping varied: the *every* target (43b) was true with a global or no DII, and false with a local DII, while the *no* and the *exactly two* targets (44b, 45b) were false with a global or no DII, and true with a local DII. Note that this means the experiment was not able to differentiate between a global DII and no DII at all.

The target *believe* sentences were identical to the target *wonder* sentences, except that *wonders whether* was replaced by *believes that*. True and False controls involved a few more replacements. Examples are given in (46).

- (46) Example of *believe* sentences. Differences with *wonder* highlighted in boldface.
- a. *Exactly two-believe*; False: Exactly two detective **believe that** the **gardener may have** committed the crime.
 - b. *Every-believe*; Target: Every detective **believes that** the **maid**, the **cook**, or the **butler** committed the crime.
 - c. *No-believe*; True: No detective **believes that** the **gardener may have** committed the crime.

5.4 Participants

For each survey version, we recruited 16 participants for each of the six possible orders for the Target, True and False sentences (hence a total of 576 HITs were posted on Mechanical Turk, paid 30ct each). Again, a few participants managed to take multiple surveys. All retakes were discarded from the analysis, and two participants who took the survey 5 or more times were not paid for their retakes and their HITs were offered to new participants. The data from 9 participants who reported native languages other than English were also discarded.

In all, after removing all retakes and non-native speakers, we had data from 543 unique native English speakers (age range: 19–74).

5.5 Results

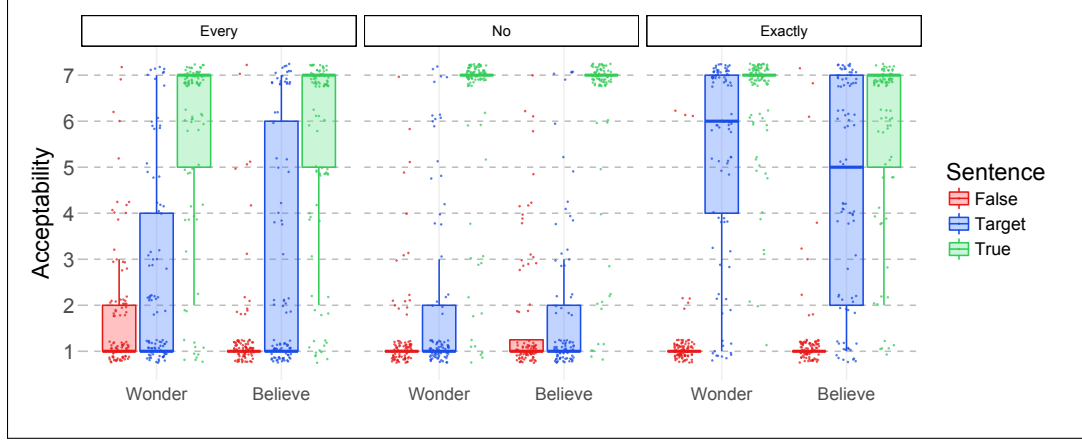


Figure 4 Experiment 2: Acceptability for each Verb, Complement, and Sentence (boxes indicate median and quartiles, each dot represents an individual answer).

The results are presented in Figure 4. Since Target sentences varied as to which readings they made true or false depending on quantifiers, we could not simply compare True and Target sentences across quantifiers. Instead of *CONDITION*, we defined two factors which had a consistent interpretation across *QUANTIFIERS*. *BASELINE* had value 0 on all False sentences and 1 on all True sentences, so it measured the maximum acceptability. The second factor, *LOCAL*, encoded the effect of local DIIs on acceptability. It had value 0 on all True and False sentences, where the truth of the local DII reading matched the truth of the literal and global DII readings and therefore had no effect. The value of the *BASELINE* and *LOCAL* predictors on Targets depended on the quantifier, as illustrated in Figure 5.

The *Every* Target was true under a global DII reading or in the absence of a DII, and false under a local DII reading. Therefore, the effect of local DIIs here was to reduce the acceptability compared to the True sentence, so the two factors were fixed at $\text{BASELINE} = 1$ and $\text{LOCAL} = -1$. As a result, the model predicted an increase in acceptability of $(\beta_{\text{BASELINE}} - \beta_{\text{LOCAL}})$ compared to False controls (on the logit scale, modulo other predictors).

No and *Exactly two* Targets were false under a global DII reading or in the absence of DII, and true under a local DII reading. Therefore, the effect of local DIIs was to increase acceptability compared to the False sentence. The values of the factors were thus fixed at $\text{BASELINE} = 0$ and $\text{LOCAL} = +1$. As a result, the

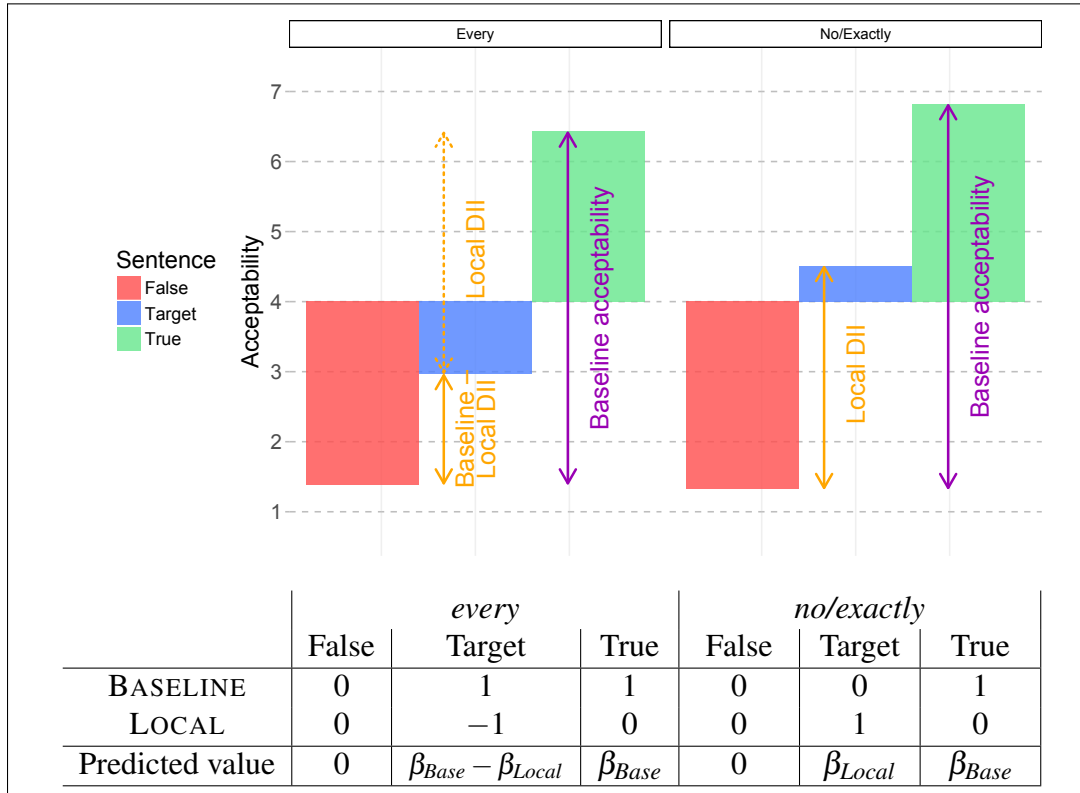


Figure 5 This figure illustrates how we detected local DIIs in the model for Experiment 2. *Every* targets are true **except** under a local DII reading, while *no* and *exactly* targets are true **only** under a local DII reading. The predicted value corresponds to the estimated position of each sentence on the logit scale (ignoring other factors).

model predicted an increase in acceptability of β_{LOCAL} compared to False controls. To sum up, the factor BASELINE always encoded the difference between True and False sentences, while LOCAL encoded the difference between True and Target for *every*, and the difference between Target and False for *no* and *exactly* two. Using these factors, we fitted a mixed-effects ordinal logistic regression to all responses, the details and results of which are described in Table 5.

We observed that LOCAL had a significant effect with baseline quantifier *every*, and that its effect was stronger with *exactly two* and weaker with *no*. Crucially, LOCAL did not interact with Verb at all (model comparison: $\chi^2(3) = 4.6, p = .21$). As a post-hoc analysis we fitted a model to *No* items only, which showed that LOCAL had a significant effect there as well, although numerically small ($\chi^2(2) = 8.7, p = .013$).

	β	z-value	p	
BASELINE	3.6	13.1	< .001	***
LOCAL	2.4	10.6	< .001	***
VERB	1.3	3.5	< .001	***
[QUANT: <i>no</i>]	-0.2	-0.9	0.390	
[QUANT: <i>exactly</i>]	-1.2	-3.7	< .001	***
BASELINE \times VERB	-1.0	-2.2	0.030	*
LOCAL \times VERB	0.4	1.1	0.280	
BASELINE \times [QUANT: <i>no</i>]	1.2	3.5	0.001	***
BASELINE \times [QUANT: <i>exactly</i>]	1.8	4.5	< .001	***
LOCAL \times [QUANT: <i>no</i>]	-1.9	-5.6	< .001	***
LOCAL \times [QUANT: <i>exactly</i>]	1.8	4.7	< .001	***
VERB \times [QUANT: <i>no</i>]	-1.8	-3.4	0.001	***
VERB \times [QUANT: <i>exactly</i>]	-1.7	-2.6	0.010	**
BASELINE \times VERB \times [QUANT: <i>no</i>]	1.4	2.0	0.049	*
BASELINE \times VERB \times [QUANT: <i>exactly</i>]	2.2	2.8	0.005	**
LOCAL \times VERB \times [QUANT: <i>no</i>]	.002	.002	0.998	
LOCAL \times VERB \times [QUANT: <i>exactly</i>]	0.5	0.7	0.467	

Table 5 Experiment 2: Full results of the statistical model. Factors of theoretical interest are highlighted for readability. Model structure: Answer \sim (Baseline + Local)*Verb*Quantifier + (1 + Baseline|Subject). The standard deviation of the random effects was .52 for the intercept and .58 for the slope. Significance levels: .05 (*), .01 (**), .001 (***)

5.6 Replications with different contexts

As with the previous experiment, one may worry that our results were affected by the specific scenario we used. This is particularly relevant in Experiment 2, since ignorance may be treated differently in the context of a murder investigation than in everyday situations. To address this issue, we carried out a follow-up experiment with two new contexts. The details of this follow-up experiment can be found in Appendix B. Overall, the results were very similar. In particular, local DIIs were present under *every*, and even more so under *exactly*. They were completely absent under *no*, unlike in the original experiment where we did detect a small effect of local DIIs under *no*. Finally, we observed no significant differences between *wonder* and *believe* in the follow-up experiment.

5.7 Discussion

In Experiment 2 we investigated the projection behavior of DIIs in quantified sentences. We observed three things: first, DIIs are sometimes computed locally in the scope of quantified subjects. Second, there was variability across quantifiers: local DIIs were more frequently computed under *exactly n* than under *every*, and less frequently, if at all, under *no*. Finally, there was no observable difference between *wonder* and *believe* in the amount of local DIIs they give rise to.

Going back to Table 1 on page 20, the first result tells us that we can rule out the first line (only global strengthening), while the second result eliminates the last line (only local strengthening). Finally, the third result indicates that there is no reason to postulate different mechanisms for the derivation of DIIs with *believe* and *wonder*.

The pattern we observed with the various quantifiers is quite similar to what has been observed in previous experimental work on local strengthening (in particular Chemla & Spector 2011, Potts et al. 2016). In these studies, the scalar item *some* was sometimes interpreted as ‘some but not all’ in the scope of *every*, more frequently so in the scope of *exactly n*, but less so in the scope of *no*. One key difference is that our overall results offer no clear evidence for local strengthening under *no*, while *some* does occasionally receive a ‘some but not all’ interpretation under *no*. Of course, there are clear differences between the design of our experiment and the sentence-picture verification task used to test local strengthening of *some*, which may be responsible for this contrast (especially given the effect of small variations in design found by van Tiel et al. 2018).

6 General discussion

6.1 Discussion of the two experiments

The two experiments discussed in the previous sections were conducted with the goal to investigate the nature of DIIs with respect to two parameters: structure-sensitivity and locality. More specifically, Experiment 1 tested the strength of DIIs with different kinds of complements. Experiment 2, on the other hand, investigated the interaction between DIIs and quantifiers with different monotonicity properties. Furthermore, since both experiments tested DIIs under *believe* and *wonder*, the results inform us as to whether we should aim for a unified account of DIIs under *believe* and *wonder*, or rather for a lexically-specific account.

The results of Experiment 1 indicate that DIIs depend on the structure of the complement. Concretely, DIIs are significantly stronger with Disjunction/Conjunction complements (e.g., *whether/that Sophie, Bill, or Mary arrived* and *which/one of Sophie, Bill, and Mary arrived*) than with NP/NumP complements (e.g., *which/one of her (three) children arrived*). These results demonstrate that the strengthening

mechanism that gives rise to DIIs is structure-sensitive. In this regard, we concur with Roelofsen & Uegaki (2016), who arrived at the same conclusion based on introspective judgments from a small group of informants (although our results suggest that the contrast between NP and NumP complements reported by Roelofsen & Uegaki 2016 is not robust across different contexts).

The results of Experiment 2 indicate that DIIs can arise locally in the scope of subject quantifiers, and furthermore that the presence of local DIIs depends on the monotonicity properties of the quantifier. Specifically, local DIIs are observed under *every*, and even more strongly under *exactly two* but they are absent under *no*, a pattern which is similar to that found in previous experimental work on the local strengthening of scalar items like *some* and *or* (Chemla & Spector 2011, Potts et al. 2016). These results point to the conclusion that DIIs arise from a strengthening mechanism that is *optionally* local. Obligatory global strengthening would not capture the DII effects under *every* and *exactly two* while obligatory local strengthening would not capture the absence of DIIs under *no*. These results are incompatible with Roelofsen & Uegaki's (2016) analysis of DIIs, which incorporates an exhaustification operator in the lexical semantics of *wonder* (predicting obligatory local strengthening).

Furthermore, in view of the general similarity between *wonder* and *believe* in our results, the most parsimonious account would be one which assumes a single strengthening mechanism that is responsible for DIIs with both *wonder* and *believe*, rather than positing verb-specific mechanisms. It should be noted that we did find a small but significant difference between *wonder* and *believe* with NP/NumP complements in Experiment 1. However, as discussed in §4.6, it is possible that this was caused by differences between the complements involved (i.e., *which of...* vs. *that one of...*), and not by the verbs themselves. Furthermore, the relevant differences were not observed in the replication of Experiment 1 (cf. Appendix A).

Taken together, the results of our experiments point to the conclusion that the best approach to DIIs is one based on a *structure-sensitive* strengthening mechanism that *optionally applies locally* (i.e. Approach 2 in Table 1 on page 20). Moreover, the most parsimonious account would be one that derives DIIs triggered by *wonder* and *believe* in a unified way.

6.2 A possible implementation

In this section, we outline an account of DIIs that captures their structure sensitivity and optional locality. Specifically, such an account of DIIs can be implemented by employing a covert exhaustivity operator, which can be optionally inserted at any sentential node in a structure, modulo pragmatic considerations (Chierchia et al. 2012). Below, we will define an appropriate exhaustivity operator within the

framework of inquisitive semantics, and illustrate how it derives DIIs with *believe* and *wonder*. Note that the goal of this section is to demonstrate that it is *possible* to have a concrete analysis of DIIs consistent with the structure-sensitivity and optional locality observed in our experimental results, rather than to argue that the specific implementation provided here is more advantageous than other accounts capturing the same characteristics of DIIs.

The exhaustivity operator EXH can be defined as in (47) within the current framework, where $\text{IE}_\Psi(\varphi)$ picks out those elements of the set of formal alternatives Ψ that are ‘innocently excludable’ (Fox 2007), as defined in (48). Here, we only define EXH with respect to a non-inquisitive prejacent for the sake of simplicity, as all relevant cases in our analysis will involve a non-inquisitive prejacent.

(47) Exhaustification

- a. $\llbracket \text{EXH}_\Psi(\varphi) \rrbracket = \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(\varphi) \text{ and} \\ \forall \psi \in \text{IE}_\Psi(\varphi) : w \notin \text{info}(\psi) \end{array} \right\} \right\}^\downarrow$
- b. $\text{presup}(\text{EXH}_\Psi(\varphi)) = \text{presup}(\varphi)$

(48) Innocently excludable formal alternatives (Fox 2007)

$$\text{IE}_\Psi(\varphi) := \left\{ \psi \in \Psi \mid \begin{array}{l} \psi \text{ is contained in every maximal set } \Psi' \subseteq \Psi \\ \text{such that } \{\neg\chi \mid \chi \in \Psi'\} \cup \{\varphi\} \text{ is consistent} \end{array} \right\}$$

The structure-sensitivity of exhaustivity is guaranteed by restricting the set of formal alternatives for exhaustification to those that are structurally simpler than the prejacent, following Katzir (2007):

(49) Structurally determined formal alternatives (Katzir 2007)

$$\text{ALT}(\varphi) := \{ \psi \mid \psi \lesssim \varphi \}$$

where $\psi \lesssim \varphi$ iff φ can be transformed into ψ by a finite series of deletions, contractions, and replacements of constituents in φ with constituents of the same category taken from the lexicon or the set of subtrees of φ .

Given this setup, we can account for the pattern of DIIs observed in our experiments. First, a non-quantified sentence with *believe* can have the following LF:

- (50) EXH_Ψ [*x* believes that Ann, Bill, or Carol did it].

This LF has the interpretation shown below, assuming that the set Ψ is restricted to the structural alternatives of the prejacent, i.e., $\text{ALT}(x \text{ believes that } A, B, \text{ or } C)$:

- (51) $\llbracket \text{EXH}_\Psi(x \text{ believes that } A, B, \text{ or } C) \rrbracket$

$$\begin{aligned}
&= \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(x \text{ believes that } A, B, \text{ or } C) \ \& \\ \forall \psi \in \text{IE}_\Psi(x \text{ believes that } A, B, \text{ or } C) : w \notin \text{info}(\psi) \end{array} \right\} \right\}^\downarrow \\
&= \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(x \text{ believes that } A, B, \text{ or } C) \\ \& w \notin \text{info}(x \text{ believes that } A \text{ or } B) \\ \& w \notin \text{info}(x \text{ believes that } B \text{ or } C) \\ \& w \notin \text{info}(x \text{ believes that } A \text{ or } C) \\ \& w \notin \text{info}(x \text{ believes that } A) \\ \& w \notin \text{info}(x \text{ believes that } B) \\ \& w \notin \text{info}(x \text{ believes that } C) \end{array} \right\} \right\}^\downarrow
\end{aligned}$$

This interpretation captures the DII, i.e., that x is ignorant about each of A , B , and C . To see this, first suppose that x believes that A in w . This is incompatible with the conjunct $w \notin \text{info}(x \text{ believes that } A)$. Now suppose that x believes not- A in w . Then, the prejacent $w \in \text{info}(x \text{ believes that } A, B \text{ or } C)$ is satisfied only if $w \in \text{info}(x \text{ believes that } B \text{ or } C)$, but this contradicts the conjunct $w \notin \text{info}(x \text{ believes that } B \text{ or } C)$. Ignorance as to B and C follows in the same manner.

Similarly, in the case of *wonder*, we have the following LF and interpretation:

$$(52) \quad \text{EXH}_\Psi [x \text{ wonders whether } A, B, \text{ or } C]$$

$$(53) \quad \llbracket \text{EXH}_\Psi(x \text{ wonders whether } A, B, \text{ or } C) \rrbracket$$

$$\begin{aligned}
&= \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(x \text{ wonders whether } A, B, \text{ or } C) \ \& \\ \forall \psi \in \text{IE}_\Psi(x \text{ wonders whether } A, B, \text{ or } C) : w \notin \text{info}(\psi) \end{array} \right\} \right\}^\downarrow \\
&= \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(x \text{ wonders whether } A, B, \text{ or } C) \\ \& w \notin \text{info}(x \text{ wonders whether } A \text{ or } B) \\ \& w \notin \text{info}(x \text{ wonders whether } B \text{ or } C) \\ \& w \notin \text{info}(x \text{ wonders whether } A \text{ or } C) \\ \& w \notin \text{info}(x \text{ wonders whether } A) \\ \& w \notin \text{info}(x \text{ wonders whether } B) \\ \& w \notin \text{info}(x \text{ wonders whether } C) \end{array} \right\} \right\}^\downarrow
\end{aligned}$$

Again, the interpretation captures the DII with respect to A , B and C . As we have done above, suppose that x believes that A in w . This is incompatible with the basic ignorance condition of the prejacent $w \in \text{info}(x \text{ wonders whether } A, B, \text{ or } C)$. Now suppose that x believes not- A in w . Then, the prejacent $w \in \text{info}(x \text{ wonders whether } A, B \text{ or } C)$ is satisfied only if $w \in \text{info}(x \text{ wonders whether } B \text{ or } C)$, but this contradicts the conjunct $w \notin \text{info}(x \text{ wonders whether } B \text{ or } C)$. Ignorance as to B and C follows in the same way.

A remark is in order regarding the one-disjunct formal alternatives in the bottom three lines in (53), e.g., *x* wonders whether A. We assume that the *whether* complements in these formal alternatives are not proper polar question complements, but rather alternative question complements involving a single disjunct. More specifically, we take it that the semantic contribution of *whether* in these cases, and in alternative questions more generally, is simply to pass up the semantic value of its sister node, as in (54).¹⁸

$$(54) \quad \llbracket \text{whether } \varphi \rrbracket = \llbracket \varphi \rrbracket$$

Given this semantic contribution of *whether*, the one-disjunct formal alternatives in (53) are interpreted as follows:

$$(55) \quad \llbracket x \text{ wonders whether } A \rrbracket = \{ \{ w \mid \text{DOX}_x^w \not\subseteq \llbracket A \rrbracket \wedge \text{INQ}_x^w \subseteq \llbracket A \rrbracket \} \}^\downarrow$$

The semantic value in (55) amounts to $\{\emptyset\}$, a contradiction. More generally, our baseline semantics for *wonder* yields a contradiction for any non-inquisitive complement (Ciardelli & Roelofsen 2015). Thus, the conditions involving one-disjunct formal alternatives in the bottom three lines in (53) are satisfied by any world *w*.

Katzir’s structure-sensitive notion of formal alternatives predicts that exhaustification gives rise to DIIs with Disjunction/Conjunction complements, but not with NP/NumP complements. This is because Disjunction/Conjunction complements always have complements with fewer disjuncts/conjuncts as their formal alternatives, as indicated in (56), while such forms do not count as formal alternatives for NP/NumP complements, as indicated in (57) (where we assume that $\llbracket \text{NP} \rrbracket = \{A, B, C\}$):

$$(56) \quad \begin{array}{lll} \text{a.} & \text{that/whether } A \text{ or } B & \lesssim \text{that/whether } A, B, \text{ or } C & \text{Disjunction} \\ \text{b.} & \text{one/which of } A \text{ and } B & \lesssim \text{one/which of } A, B, \text{ and } C & \text{Conjunction} \end{array}$$

$$(57) \quad \begin{array}{lll} \text{a.} & \text{one/which of } A \text{ and } B & \not\lesssim \text{one/which of the NP} & \text{NP} \\ \text{b.} & \text{one/which of } A \text{ and } B & \not\lesssim \text{one/which of the three NPs} & \text{NumP} \end{array}$$

This said, we should note that NP/NumP complements may have ‘subdomain alternatives’ equivalent to the formal alternatives with fewer disjuncts/conjuncts in (57) (Chierchia 2013, Roelofsen & Uegaki 2016). We take the occasional DIIs

¹⁸ It is widely assumed that polar questions do not amount to alternative questions with a single disjunct (Karttunen 1977, Biezma & Rawlins 2012, Guerzoni & Sharvit 2014, Roelofsen 2015). This is reflected, for instance, by the fact that they involve different prosody (Pruitt & Roelofsen 2013) and have different patterns for NPI licensing (Guerzoni & Sharvit 2014). As a consequence, the polar question ‘whether A (or not)’ does not qualify as a formal alternative to the alternative question ‘whether A or B or C’ in Katzir’s theory.

observed in the NP/NumP conditions in Experiment 1 to be due to the availability of such subdomain alternatives.¹⁹ The contrast in the amount of DIIs between the Disjunction/Conjunction conditions and the NP/NumP conditions is thus ultimately explained in terms of the difference in activation between structural alternatives and subdomain alternatives (see e.g., [Denić & Chemla 2018](#), for a similar contrast between disjunctions and indefinites).

Moreover, given that EXH can be inserted locally, the possibility of local DIIs within the scope of a quantified subject in Experiment 2 can be accounted for as well. Concretely, the local DIIs under *every detective* are derived on the basis of the following LFs:

- (58) a. [Every detective]₁ 1 EXH_Ψ [*x*₁ believes that A, B, or C].
 b. [Every detective]₁ 1 EXH_Ψ [*x*₁ wonders whether A, B, or C].

In contrast to the local application of EXH, the matrix application of EXH, as in the following LFs, would derive the global DIIs:

- (59) a. EXH_Ψ [[Every detective]₁ 1 [*x*₁ believes that A, B, or C]].
 b. EXH_Ψ [[Every detective]₁ 1 [*x*₁ wonders whether A, B, or C]].

While the local DII readings derived from (58) are straightforward, it is useful to spell out the precise predictions that are made concerning global DIIs. The LF in (59a) is interpreted as follows:

$$(60) \quad \llbracket (59a) \rrbracket = \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(\text{every detective believes that A, B, or C}) \\ \& w \notin \text{info}(\text{every detective believes that A or B}) \\ \& w \notin \text{info}(\text{every detective believes that B or C}) \\ \& w \notin \text{info}(\text{every detective believes that A or C}) \\ \& w \notin \text{info}(\text{every detective believes that A}) \\ \& w \notin \text{info}(\text{every detective believes that B}) \\ \& w \notin \text{info}(\text{every detective believes that C}) \end{array} \right\} \right\}^{\downarrow}$$

The exhausted meaning in (60) entails that none of the three disjuncts are believed by every detective and that none of the three disjuncts is excluded by every detective. In other words, for each disjuncts there is at least one detective who is not convinced that it is true and one detective (possibly the same) who is not convinced that it is false.

¹⁹ Note that we observed differences between the different versions of Experiment 1 with respect to the NumP condition (some DIIs with *wonder* in the original experiment, some with both verbs in the Geography context, and as much as with disjunctions in the Bus context). This would suggest that the activation of subdomain alternatives depends on various factors (for instance, in the bus context the NP/NumP is in object position), but identifying these factors is beyond the scope of this paper.

The LF in (59b) is interpreted as follows:²⁰

$$(61) \quad \llbracket (59b) \rrbracket = \left\{ \left\{ w \mid \begin{array}{l} w \in \text{info}(\text{every detective wonders whether A, B, or C}) \\ \& w \notin \text{info}(\text{every detective wonders whether A or B}) \\ \& w \notin \text{info}(\text{every detective wonders whether B or C}) \\ \& w \notin \text{info}(\text{every detective wonders whether A or C}) \end{array} \right\} \right\} \downarrow$$

This interpretation entails that there is a detective who is ignorant about A. This follows from the following reasoning: first, given the ignorance condition of the prejacent, no detective believes A. Furthermore, if a detective believes not-A, then the detective should entertain B or C, given the entertainment condition of the prejacent. However, the third line in (61) states that this does not hold for at least one detective. The same holds for B and C as well. Overall, (61) entails that each of A, B or C is considered possible by at least one detective, and none is believed by any detective.

Chierchia et al. (2012) discuss two possible principles for disambiguating between the different parses of a sentence corresponding to different placements of EXH. First, they consider a preference for the strongest possible parse (cf. Dalrymple et al. 1998). This would predict a preference for local exhaustification under *every* and global exhaustification with *no*, but it wouldn't make any prediction for *exactly* *n* with $n > 1$ since the local and global parses are logically independent in this case (and both stronger than the parse without exhaustification). This principle is also problematic in that it would require a number of comparisons which increases exponentially with the number of possible placements of EXH. Chierchia et al. (2012) then consider an alternative principle which is less demanding in terms of the required number of comparisons: a parse with EXH is marked if it is weaker than the corresponding parse without EXH, but different parses with EXH are not compared to each other. This makes parses with EXH in downward-entailing environments marked, but does not distinguish between the global and local parses for *every* and *exactly*. More recently, Bergen et al. (2016) proposed a probabilistic pragmatic model for the disambiguation of sentences with embedded scalar terms, and Potts et al. (2016) showed that the quantitative predictions of such a model regarding the rates of local/global readings can be rather accurate. Most importantly, the pattern we observed is similar to their experimental results on embedded scalars, and any refinement of the theory of local exhaustification that explains the robust local effect under non-monotonic quantifiers and the markedness of exhaustification under *no* would explain the pattern we observed with DIIs as well.

²⁰ Here we omit one-disjunct formal alternatives because they are inconsequential, as discussed above.

7 Concluding remarks

In this paper, we have presented two experiments concerning *distributive ignorance inferences* (DIIs) triggered by sentences involving the attitude predicates *wonder* and *believe*. The results suggest that such DIIs result from a strengthening mechanism which is sensitive to the syntactic structure of the complement of the verbs involved (cf. Katzir 2007) and which optionally applies locally, as part of the semantic composition process (cf. Chierchia et al. 2012).

Looking beyond the domain of DIIs with *wonder* and *believe*, our experimental results make a contribution to the experimental investigation of local strengthening more generally. Of particular interest is the finding that the extent to which DIIs arise locally, i.e., in the scope of quantifiers with different monotonicity properties, is similar to the extent to which scalar items like *some* and *or* are strengthened locally. An interesting avenue for future research would be to consider whether this finding, as well as further experimental work on DIIs, may be able to tease apart different theoretical accounts of local strengthening (Chierchia et al. 2012, Geurts & van Tiel 2013, Potts et al. 2016). In particular, the approach of Chierchia et al. (2012), which derives local strengthening by postulating covert exhaustivity operators, seems to be able to derive DIIs quite straightforwardly. The approach of Bergen et al. (2016) and Potts et al. (2016) is framed in terms of lexical underspecification rather than presence/absence of a covert operator, but it could easily be reframed in term of exhaustification. For the pragmatic approach of Geurts & van Tiel (2013), which relies on typicality effects in the interpretation of specific lexical items and narrow focus on scalar terms, it is, at least at first sight, not so clear how the account of scalar items like *some* and *or* could be extended. The kind of inferences discussed here do not seem to arise from any specific lexical item, but rather from the interplay between several elements, making it very unlikely that the implicature would be conventionalized, and it is not clear that local DIIs require any specific focus. However, a careful discussion of this issue must be left for another occasion.

Another avenue for further research is to compare the proposal made here with that of Blumberg (2017) (which was published after the present paper had been submitted for publication). Our experimental findings on local DIIs seem problematic for Blumberg's proposal, but may be compatible with a refined version of it. Blumberg points out that the two proposals also make different predictions about DIIs arising when *wonder* takes a polar disjunctive question as its complement, rather than an alternative question. Since the judgments in these cases seem rather subtle, a proper comparison of the two approaches in this regard would have to be based on further experimental work.

References

- Alonso-Ovalle, Luis. 2006. *Disjunction in alternative semantics*: University of Massachusetts, Amherst dissertation.
- AnderBois, Scott. 2012. Focus and un informativity in Yukatek Maya questions. *Natural Language Semantics* 20. 349–390. <https://doi.org/10.1007/s11050-012-9084-3>.
- Bartels, Christine. 1999. *The intonation of English statements and questions: a compositional interpretation*. New York: Routledge.
- Benz, Anton & Nicole Gotzner. 2014. Embedded implicatures revisited: issues with the truth-value judgment paradigm. In *Formal & experimental pragmatics workshop*, vol. 9 20, 1–6. <https://doi.org/10.3765/sp.9.20>.
- Bergen, Leon, Roger Levy & Noah Goodman. 2016. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics* 9(20). 1–83. <https://doi.org/10.3765/sp.9.20>.
- Biezma, Maria & Kyle Rawlins. 2012. Responding to alternative and polar questions. *Linguistics and Philosophy* 35(5). 361–406. <https://doi.org/10.1007/s10988-012-9123-z>.
- Blumberg, Kyle. 2017. Ignorance implicatures and non-doxastic attitude verbs. In Alexandre Cremers, Thom van Gessel & Floris Roelofsen (eds.), *The 21st Amsterdam colloquium*, 135–144.
- Champollion, Lucas, Ivano Ciardelli & Floris Roelofsen. 2015. Some questions in typed inquisitive semantics. Talk at workshop on Questions in Logic and Semantics, Amsterdam.
- Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28. 359–400. <https://doi.org/10.1093/jos/ffq023>.
- Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and beyond*, Oxford: Oxford University Press.
- Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press.
- Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2009. Hurford’s constraint and the theory of scalar implicatures: evidence for embedded implicatures. In Paul Egré & Giorgio Magri (eds.), *Presuppositions and implicatures*, 47–62. Cambridge, MA: MIT Press.
- Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. Scalar implicature as a grammatical phenomenon. In Paul Portner, Claudia Maienborn & Klaus von Stechow (eds.), *Semantics: An international handbook of natural language meaning*, vol. 3, 2297–2331. Berlin: Mouton de Gruyter.

- Christensen, Rune Haubo B. 2015. ordinal—regression models for ordinal data. R package version 2015.6-28. <http://www.cran.r-project.org/package=ordinal/>.
- Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2012. Inquisitive semantics. NASSLLI lecture notes, available via projects.illc.uva.nl/inquisitivesemantics.
- Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2013. Inquisitive semantics: A new notion of meaning. *Language and Linguistics Compass* 7(9). 459–476. <https://doi.org/10.1111/lnc3.12037>.
- Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2015. Inquisitive semantics. European Summer School in Logic, Language and Information lecture notes, available via projects.illc.uva.nl/inquisitivesemantics.
- Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2018. *Inquisitive semantics*. Oxford University Press. To appear.
- Ciardelli, Ivano & Floris Roelofsen. 2015. Inquisitive dynamic epistemic logic. *Synthese* 192(6). 1643–1687. <https://doi.org/10.1007/s11229-014-0404-7>.
- Ciardelli, Ivano & Floris Roelofsen. 2017. Hurford’s constraint, the semantics of disjunctions, and the nature of alternatives. *Natural Language Semantics* 25(3). 199–222. <https://doi.org/10.1007/s11050-017-9134-y>.
- Ciardelli, Ivano & Floris Roelofsen. 2018. An inquisitive perspective on modals and quantifiers. *Annual Review of Linguistics* 4. 129–149. <https://doi.org/10.1146/annurev-linguistics-011817-045626>.
- Clifton, Charles & Chad Dube. 2010. Embedded implicatures observed: a comment on Geurts and Pouscoulous (2009). *Semantics and Pragmatics* 3(7). 1–13. <https://doi.org/10.3765/sp.3.7>.
- Cohen, Michael. 2017. Neg-raising and question embedding. Presented at the UCSC-Stanford Workshop on Sentence Types.
- Coppock, Elizabeth & Thomas Brochhagen. 2013. Diagnosing truth, interactive sincerity, and depictive sincerity. In *Semantics and Linguistic Theory (SALT)* 23, 358–375. <https://doi.org/10.3765/salt.v23i0.2662>.
- Csipak, Eva & Sarah Zobel. 2014. A condition on the distribution of discourse particles across types of questions. In *North East Linguistic Society (NELS)* 44, 83–94.
- Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam McHombo & Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2). 159–210. <https://doi.org/10.1023/A:1005330227480>.
- Denić, Milica & Emmanuel Chemla. 2018. Q-spreading in child language as distributive inferences. Ms. LSCP & IJN. <https://semanticsarchive.net/Archive/GEzMDU4N/Denic-Chemla-quantifier-spreading-distributive-inferences.html>.
- Fox, Danny. 2007. Free choice disjunction and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in*

- compositional semantics*, 71–120. Basingstoke: Palgrave Macmillan.
- Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19. 87–107. <https://doi.org/10.1007/s11050-010-9065-3>.
- Gajewski, Jon Robert. 2007. Neg-raising and polarity. *Linguistics and Philosophy* 30(3). 289–328. <https://doi.org/10.1007/s10988-007-9020-z>.
- Geurts, Bart. 2009. Scalar implicature and local pragmatics. *Mind and Language* 24. 51–79. <https://doi.org/10.1111/j.1468-0017.2008.01353.x>.
- Geurts, Bart & Nausicaa Pouscoulous. 2009. Embedded implicatures?!? *Semantics and Pragmatics* 2(4). 1–34. <https://doi.org/10.3765/sp.2.4>.
- Geurts, Bart & Bob van Tiel. 2013. Embedded scalars. *Semantics and Pragmatics* 6(9). 1–37. <https://doi.org/10.3765/sp.6.9>.
- Geurts, Bart & Bob van Tiel. 2016. When “all the five circles” are four: new exercises in domain restriction. *Topoi* 1(35). 109–122. <https://doi.org/10.1007/s11245-014-9293-0>.
- Guerzoni, Elena & Yael Sharvit. 2007. A question of strength: on NPIs in interrogative clauses. *Linguistics and Philosophy* 30(3). 361–391. <https://doi.org/10.1007/s10988-007-9014-x>.
- Guerzoni, Elena & Yael Sharvit. 2014. Whether or not anything but not whether anything or not. In Luka Crnic & Uli Sauerland (eds.), *The art and craft of semantics: A festschrift for Irene Heim*, MIT Working Papers in Linguistics.
- Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10(1). 41–53.
- Hintikka, Jaakko. 1962. *Knowledge and belief: an introduction to the logic of the two notions*. Cornell University Press.
- Horn, Laurence R. 2006. The border wars: A neo-Gricean perspective. *Where semantics meets pragmatics* 16. 21–48.
- Karttunen, L. 1974. Presupposition and linguistic context. *Theoretical linguistics* 1(1–3). 181–194. <https://doi.org/10.1515/thli.1974.1.1-3.181>.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3–44. <https://doi.org/10.1007/bf00351935>.
- Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6). 669–690. <https://doi.org/10.1007/s10988-008-9029-y>.
- Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In Yukio Otsu (ed.), *The third Tokyo conference on psycholinguistics*, 1–25.
- Križ, Manuel. 2015. *Aspects of homogeneity in the semantics of natural language*: University of Vienna dissertation.
- Mayr, Clemens. 2017. Predicting polar question embedding. In *Proceedings of Sinn und Bedeutung 21*, .

- Onea, Edgar & Markus Steinbach. 2012. Where question, conditionals and topics converge. In Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit Weidmann-Sassoon, Katrin Schulz & Matthijs Westera (eds.), *Logic, language, and meaning. selected papers from the 18th Amsterdam colloquium*, 42–51. Berlin Heidelberg: Springer.
- Phillips, Jonathan & B R George. 2018. Knowledge wh and false beliefs: Experimental investigations. *Journal of Semantics* ffy004. <https://doi.org/10.1093/semant/ffy004>.
- Potts, Christopher, Daniel Lassiter, Roger Levy & Michael C Frank. 2016. Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics* 33(4). 755–802. <https://doi.org/10.1093/jos/ffv012>.
- Pruitt, Kathryn & Floris Roelofsen. 2013. The interpretation of prosody in disjunctive questions. *Linguistic Inquiry* 44. 632–650. https://doi.org/10.1162/LING_a_00141.
- R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <http://www.R-project.org/>.
- Roelofsen, Floris. 2015. The semantics of declarative and interrogative lists. Manuscript, ILLC, University of Amsterdam.
- Roelofsen, Floris, Michele Herbstritt & Maria Aloni. 2016. The **whether* puzzle. To appear in *Questions in Discourse*, edited by Klaus von Heusinger, Edgar Onea, and Malte Zimmermann.
- Roelofsen, Floris & Wataru Uegaki. 2016. The distributive ignorance puzzle. In *Sinn und Bedeutung* 21, .
- Rojas-Esponda, Tania. 2014. A QUD account of German *doch*. In Urtzi Etxeberria, Anamaria Fălăuş, Aritz Irurtzun & Bryan Leferman (eds.), *Sinn und Bedeutung* 18, 359–376.
- Romoli, Jacopo. 2013. A scalar implicature-based approach to neg-raising. *Linguistics and philosophy* 36(4). 291–353. <https://doi.org/10.1007/s10988-013-9136-2>.
- Starr, William B. 2014. What ‘if’? *Philosophers’ Imprint* 14(10). 1–27.
- Theiler, Nadine, Floris Roelofsen & Maria Aloni. 2017a. Deriving selectional restrictions of clause-embedding predicates. Manuscript, ILLC, University of Amsterdam.
- Theiler, Nadine, Floris Roelofsen & Maria Aloni. 2017b. What’s wrong with believing whether? In *Semantics and Linguistic Theory (SALT)* 27, 248–265. <https://doi.org/10.3765/salt.v27i0.4125>.
- Theiler, Nadine, Floris Roelofsen & Maria Aloni. 2018. A uniform semantics for declarative and interrogative complements. *Journal of Semantics* 35(3). 409–466. <https://doi.org/10.1093/jos/ffy003>.

- van Tiel, Bob, Ira Noveck & Mikhail Kissine. 2018. Reasoning with ‘some’. *Journal of Semantics* Advanced Access. <https://doi.org/10.1093/jos/ffy012>.
- Uegaki, Wataru. 2015. *Interpreting questions under attitudes*: Massachusetts Institute of Technology dissertation.

A Replication of Experiment 1 with other contexts

We replicated Experiment 1 with two other contexts, which we will call the Bus context and the Geography context.

- (62) **Bus context:** Bill is visiting his friend Sue. She will still be at work when he lands, so she cannot come to pick him up at the airport. Bill wants to spend as little as possible, so he won't take a taxi. Fortunately three bus lines run from the airport to Sue's neighborhood: line 6, line 11, and line 2E. Line 2E is an express line, and it's more expensive. Sue can pick up Bill at the bus stop on her way back from work, but she needs to know which bus line he is taking because they each serve a different stop. Knowing Bill, Sue knows that he won't take a taxi, and she is even convinced that he won't use the express line because of the extra fee. However, Sue doesn't know whether he will take line 6 or line 11.
- (63) **Geography context:** Sue, an 11th-grade student, has to choose which subjects she will take next year. Her choice depends on who will be teaching the subjects. The school has three geography teachers: Mr. Smith, Ms. Adams, and Mr. Brown. Sue knows that the school will not hire any new geography teacher next year, so one of the current teachers will have to teach 12th-grade geography. She also thinks that Ms. Adams won't be teaching next year because she will be on an exchange.

All items for each context are given in Tables 6 and 7.

	COMP	CONDITION	Sentence
<i>wonder</i>	DISJ	False	Sue wonders whether Bill is taking line 2E.
		Target	Sue wonders whether Bill is taking line 6, 11, or 2E.
		True	Sue wonders whether Bill is taking line 6 or 11.
	CONJ	False	Sue wonders whether Bill is taking line 2E.
		Target	Sue wonders which of line 6, 11 and 2E Bill is taking.
		True	Sue wonders which of lines 6 and 11 Bill is taking.
	NP	False	Sue wonders whether Bill is taking line 2E.
		Target	Sue wonders which bus line Bill is taking.
		True	Sue wonders which regular line Bill is taking.
	NUMP	False	Sue wonders whether Bill is taking line 2E.
		Target	Sue wonders which of the three lines Bill is taking.
		True	Sue wonders which of the two regular lines Bill is taking.
<i>believe</i>	DISJ	False	Sue believes that Bill is taking line 2E.
		Target	Sue believes that Bill is taking line 6, 11, or 2E.
		True	Sue believes that Bill is taking line 6 or 11.
	CONJ	False	Sue believes that Bill is taking line 2E.
		Target	Sue believes that Bill is taking one of line 6, 11 and 2E.
		True	Sue believes that Bill is taking one of lines 6 and 11.
	NP	False	Sue believes that Bill is taking line 2E.
		Target	Sue believes that Bill is taking one of the bus lines.
		True	Sue believes that Bill is taking one of the regular lines.
	NUMP	False	Sue believes that Bill is taking line 2E.
		Target	Sue believes that Bill is taking one of line 6, 11 and 2E.
		True	Sue believes that Bill is taking one of the two regular lines.
Fillers	False		Bill will probably take a taxi from the airport.
	True		Sue doesn't know at which bus stop she needs to pick up Bill.

Table 6 Stimuli for the Bus context replication of Experiment 1

Distributive ignorance inferences with *wonder* and *believe*

	COMP	CONDITION	Sentence
<i>wonder</i>	DISJ	False	Sue wonders whether or not Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue wonders whether Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Sue wonders whether Mr. Smith or Mr. Brown will teach the 12th-grade geography class next year.
	CONJ	False	Sue wonders whether or not Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue wonders which of Mr. Smith, Ms. Adams, and Mr. Brown will teach the 12th-grade geography class next year.
		True	Sue wonders which of Mr. Smith and Mr. Brown will teach the 12th-grade geography class next year.
	NP	False	Sue wonders whether or not Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue wonders which of the geography teachers will teach the 12th-grade geography class next year.
		True	Sue wonders which of the male geography teachers will teach the 12th-grade geography class next year.
	NUMP	False	Sue wonders whether or not Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue wonders which of the three geography teachers will teach the 12th-grade geography class next year.
		True	Sue wonders which of the two male geography teachers will teach the 12th-grade geography class next year.
<i>believe</i>	DISJ	False	Sue believes that Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue believes that Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Sue believes that Mr. Smith or Mr. Brown will teach the 12th-grade geography class next year.
	CONJ	False	Sue believes that Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue believes that one of Mr. Smith, Ms. Adams, and Mr. Brown will teach the 12th-grade geography class next year.
		True	Sue believes that one of Mr. Smith and Mr. Brown will teach the 12th-grade geography class next year.
	NP	False	Sue believes that Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue believes that one of the current teachers will teach the 12th-grade geography class next year.
		True	Sue believes that one of the male geography teachers will teach the 12th-grade geography class next year.
	NUMP	False	Sue believes that Ms. Adams will teach the 12th-grade geography class next year.
		Target	Sue believes that one of the three current geography teachers will teach the 12th-grade geography class next year.
		True	Sue believes that one of the two male geography teachers will teach the 12th-grade geography class next year.
Fillers	False		Sue thinks that the school will hire a new geography teacher.
	True		Sue is considering taking geography in 12th grade.

Table 7 Stimuli for the Geography context replication of Experiment 1**A.1 Design**

The design was nearly identical to Experiment 1, except that each participant saw eight items: a target, a true control and a false control for each verb, plus two fillers. The VERB factor was therefore within-participant, while the COMPLEMENT

factor remained between-participants. We further introduced a between-participants CONTEXT factor with two levels (Bus and Geography). Apart from addressing some of the reviewers’ concerns regarding Experiment 1, increasing the number of items per participant also made the survey more attractive on Mechanical Turk (for the same hourly rate, participants usually prefer longer tasks on MTurk, and some participants filter tasks below 50¢). This made recruitment much faster and reduced the risk that participants would take the survey multiple times.

Moving from 3 to 8 items means we went from 6 possible orders to 40320, so we could not test them all. We first restricted ourselves to the permutations which had the following structure: three pairs of one true and one false control or filler each, interspersed with the two targets. Furthermore, we required that the items in each pair of controls/fillers would not be controls for the same verb, or both be fillers. We had 8 sets of items ($2 \text{ Contexts} \times 4 \text{ Complements}$), and we tested each of them with 48 different orders among the 192 options satisfying our constraints.²¹

A.2 Participants

384 participants (48 for each set of items) were recruited on Amazon’s Mechanical Turk and paid 57¢ each. No two participants saw the same set of items in the same order. The data from 8 participants who reported native languages other than English were discarded. This time, we only observed three retakes, which were discarded from the analysis and not paid.

²¹ More concretely: we first picked a permutation σ in \mathfrak{S}_3 for the two true controls and the true filler (6 possibilities). We then picked σ' for the false controls and filler as one of the two permutations in \mathfrak{S}_3 such that $\forall i, \sigma'(i) \neq \sigma(i)$ (2 possibilities). Three pairs $i \in \{1, 2, 3\}$ were formed with True control/filler $\sigma(i)$ and False control/filler $\sigma'(i)$. We then picked three permutations $\sigma_1, \sigma_2, \sigma_3 \in \mathfrak{S}_2$, to randomize the order of the true and false item in each pair of controls/fillers (2^3 possibilities), and one last permutation σ_5 to randomize the order of the *wonder* and *believe* targets (2 possibilities). In all, this lead to $6 \times 2 \times 8 \times 2 = 192$ possible orders. For each set of items, we tested all 48 combinations of $\sigma, \sigma_1, \sigma_2, \sigma_3$, each with one randomly picked combination of σ' and σ_5 among the four possible.

A.3 Results

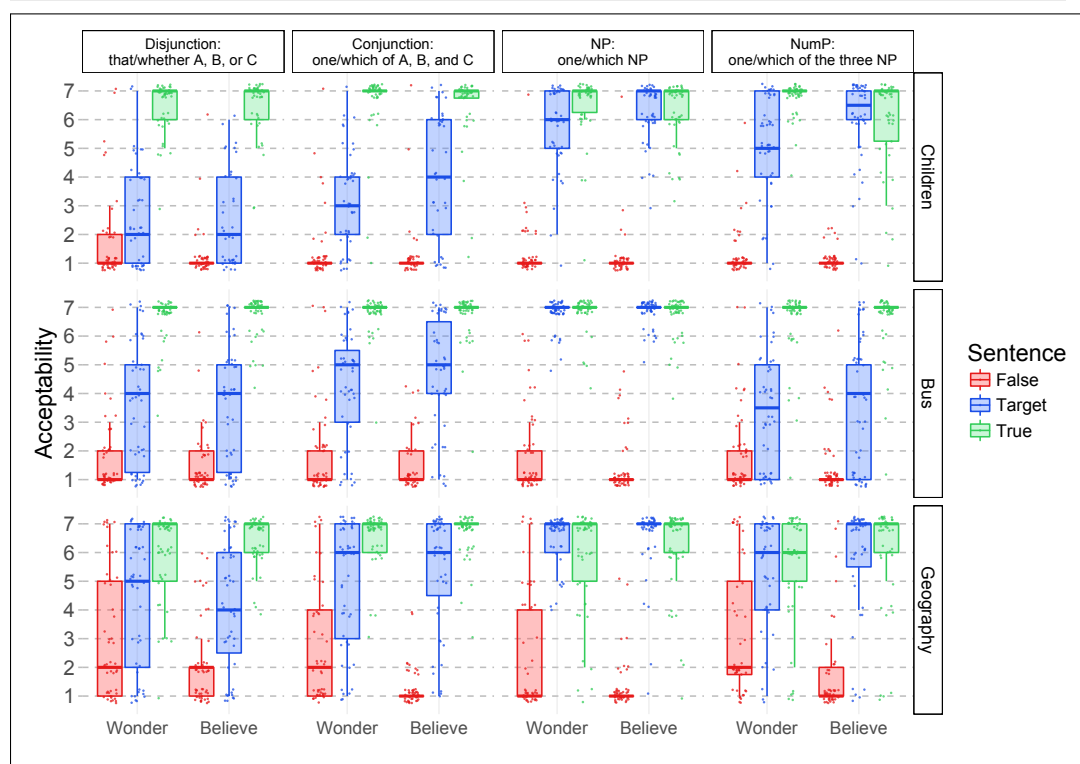


Figure 6 Replications of Experiment 1: Acceptability for each Verb, Complement, and Sentence by Context (boxes indicate median and quartiles, each dot represents an individual answer). The results of the original experiment (Children context) are repeated in the first row.

The results of the replications are presented in Figure 6. To compare the results of the replications with the original experiment, we added a *CONTEXT* factor to the mixed-model and ran it on the data from all three versions of the experiment. We also included a random slope for *VERB* since this factor varied within-participant for a subset of the participants.

- (64) `clmm(Answer~Condition*Verb*Comp*Context+(1+Verb|Subject))`
 where *CONDITION* is 0 for True controls and 1 for Targets, *VERB* is sum-coded, *COMPLEMENT* is treatment-coded with Disjunction as the baseline, and *CONTEXT* is treatment-coded, with Children as the baseline (the original experiment).

The results are presented in Table 8. We observed four significant effects of context (before any correction): targets were more acceptable in the Geography context (suggesting weaker DII), but still clearly distinct from true controls with Disjunction COMPLEMENTS, and there were some differences with NumP COMPLEMENTS. In both new contexts, NumP targets were more degraded than in the original experiment, and in the Bus context the effect of VERB was not observed.

As a post-hoc analysis, we tested the difference between the NP and NumP COMPLEMENTS by comparing a model that distinguishes the two from one that doesn't. We found a clear difference in both Bus ($\chi^2(4) = 151, p < .001$) and Geography contexts ($\chi^2(4) = 20, p < .001$), while there was none in Experiment 1.

Finally, we compared two models for the responses to targets other than NumP across contexts: one in which all interactions between CONTEXT, VERB and COMPLEMENT were included, and one without any interaction between CONTEXT and the two other factors. The model with all interactions did not show any significant improvement compared to the simpler model ($\chi^2(10) = 11.4, p = .33$), suggesting that only the NumP targets were significantly affected by context.

	β	z-value	p-value
SENTENCE:Target	-4.26	-12.99	< .001***
VERB	-0.04	-0.09	.929
COMPLEMENT:Conjunction	0.53	1.39	.165
COMPLEMENT:NP	-0.04	-0.13	.899
COMPLEMENT:NumP	0.03	0.09	.930
CONTEXT:Bus	0.65	1.70	.088
CONTEXT:Geography	-0.60	-1.74	.083
[SENTENCE:Target] × VERB	0.27	0.44	.660
[SENTENCE:Target] × [COMPLEMENT:Conjunction]	0.58	1.26	.208
[SENTENCE:Target] × [COMPLEMENT:NP]	3.44	7.76	< .001***
[SENTENCE:Target] × [COMPLEMENT:NumP]	2.94	6.57	< .001***
VERB × [COMPLEMENT:Conjunction]	0.36	0.47	.639
VERB × [COMPLEMENT:NP]	0.78	1.12	.264
VERB × [COMPLEMENT:NumP]	1.60	2.23	.026*
[SENTENCE:Target] × [CONTEXT:Bus]	0.36	0.80	.425
[SENTENCE:Target] × [CONTEXT:Geo]	2.25	5.32	< .001***
VERB × [CONTEXT:Bus]	0.20	0.28	.783
VERB × [CONTEXT:Geo]	-0.57	-0.88	.376
[COMPLEMENT:Conjunction] × [CONTEXT:Bus]	-0.18	-0.31	.758
[COMPLEMENT:NP] × [CONTEXT:Bus]	0.35	0.64	.522
[COMPLEMENT:NumP] × [CONTEXT:Bus]	-0.20	-0.38	.704
[COMPLEMENT:Conjunction] × [CONTEXT:Geo]	0.56	1.06	.289
[COMPLEMENT:NP] × [CONTEXT:Geo]	0.16	0.32	.748
[COMPLEMENT:NumP] × [CONTEXT:Geo]	-0.25	-0.51	.612
[SENTENCE:Target] × VERB × [COMPLEMENT:Conjunction]	-1.45	-1.57	.116
[SENTENCE:Target] × VERB × [COMPLEMENT:NP]	-1.77	-2.04	.041*
[SENTENCE:Target] × VERB × [COMPLEMENT:NumP]	-2.92	-3.31	< .001***
[SENTENCE:Target] × VERB × [CONTEXT:Bus]	-0.44	-0.50	.618
[SENTENCE:Target] × VERB × [CONTEXT:Geo]	0.77	0.91	.360
[SENTENCE:Target] × [COMPLEMENT:Conjunction] × [CONTEXT:Bus]	-0.11	-0.17	.862
[SENTENCE:Target] × [COMPLEMENT:NP] × [CONTEXT:Bus]	0.49	0.72	.469
[SENTENCE:Target] × [COMPLEMENT:NumP] × [CONTEXT:Bus]	-2.93	-4.59	< .001***
[SENTENCE:Target] × [COMPLEMENT:Conjunction] × [CONTEXT:Geo]	-0.64	-1.03	.304
[SENTENCE:Target] × [COMPLEMENT:NP] × [CONTEXT:Geo]	-0.53	-0.86	.392
[SENTENCE:Target] × [COMPLEMENT:NumP] × [CONTEXT:Geo]	-1.35	-2.24	.025*
VERB × [COMPLEMENT:Conjunction] × [CONTEXT:Bus]	-0.30	-0.27	.784
VERB × [COMPLEMENT:NP] × [CONTEXT:Bus]	-0.42	-0.40	.687
VERB × [COMPLEMENT:NumP] × [CONTEXT:Bus]	-1.89	-1.84	.065
VERB × [COMPLEMENT:Conjunction] × [CONTEXT:Geo]	-0.01	-0.01	.990
VERB × [COMPLEMENT:NP] × [CONTEXT:Geo]	-0.58	-0.63	.528
VERB × [COMPLEMENT:NumP] × [CONTEXT:Geo]	-1.72	-1.84	.066
[SENTENCE:Target] × VERB × [COMPLEMENT:Conjunction] × [CONTEXT:Bus]	0.95	0.73	.467
[SENTENCE:Target] × VERB × [COMPLEMENT:NP] × [CONTEXT:Bus]	1.42	1.06	.289
[SENTENCE:Target] × VERB × [COMPLEMENT:NumP] × [CONTEXT:Bus]	3.16	2.50	.012*
[SENTENCE:Target] × VERB × [COMPLEMENT:Conjunction] × [CONTEXT:Geo]	0.13	0.11	.916
[SENTENCE:Target] × VERB × [COMPLEMENT:NP] × [CONTEXT:Geo]	0.76	0.62	.535
[SENTENCE:Target] × VERB × [COMPLEMENT:NumP] × [CONTEXT:Geo]	1.76	1.46	.144

Table 8 Results of the model comparing the replications of Experiment 1 to the original experiment. Effects of interest in the original experiment are highlighted in blue, while their interaction with context are highlighted in brown. The standard deviation of the random effects was .78 for the intercept and .15 for the slope.

A short discussion of these results is presented in §4.5.

B Replication of Experiment 2 with other contexts

We replicated Experiment 2 with two other contexts, which we will call the Chefs context and the Geography context (the latter was very similar to the Geography context used to replicate Experiment 1). The contexts shared the same structure as the Detective context used in our original Experiment 2. In particular, they came in two versions: one for the quantifiers *every* and *no*, and one for the quantifier *exactly*, and they made the local DII reading true for *every*, and false for *exactly* and *no*.

(65) Chefs context:

- a. ***every/no* version:** There are three chefs taking part in a yearly evaluation, in which they can be rated 0, 1, 2 or 3 stars. They have to carry out three difficult assignments. For each assignment that they carry out in an excellent way, they get one star. They are currently waiting for the jury's final announcement. Two chefs think that they performed perfectly on one of tasks, but that they failed one of the assignments. This means that they could end up with either one or two stars. The last chef thinks that she failed to complete two of the assignments. This means that she could end up with at most one star.
- b. ***exactly* version:** There are three chefs taking part in a yearly evaluation, in which they can be rated 0, 1, 2 or 3 stars. They have to carry out three difficult assignments. For each assignment that they carry out in an excellent way, they get one star. They are currently waiting for the jury's announcement. One chef thinks that she failed to complete two of the assignments. This means that she could end up with at most one star. The two other chefs think that they failed to complete one of the assignments, so they could end up with at most two stars.

(66) Geography context:

- a. ***every/no* version:** Thirty 11th-grade students have to choose which subjects they will take next year. The students' choices depend on who will be teaching the subjects. The school has four geography teachers: Mr. Smith, Ms. Jones, Ms. Adams, and Mr. Brown. Only one of them will teach 12th-grade geography next year, and no new teacher will be recruited. All students know that Ms. Jones won't be teaching next year because she will be on an exchange. One student additionally knows that Mr. Smith won't be teaching the 12th-grade geography next year because he will be teaching the 10th-grade geography class. Two other students think that Ms. Adams won't be teaching next year. All students believe that it is possible that Mr. Brown will teach next year, and no student is aware of what others know or think.

- b. ***exactly version***: Ten 11th-grade students have to choose which subjects they will take next year. The students' choices depend on who will be teaching the subjects. The school has four geography teachers: Mr. Smith, Ms. Jones, Ms. Adams, and Mr. Brown. Only one of them will teach 12th-grade geography next year, and no new teacher will be recruited. All students know that Ms. Jones won't be teaching next year because she will be on an exchange. Two students additionally think that Mr. Smith won't be teaching the 12th-grade geography next year because he will be teaching the 10th-grade geography class. All students believe that it is possible that Mr. Brown or Ms. Adams will teach next year, and no student is aware of what others know or think.

All stimuli for the Chefs context and the Geography context are presented in Tables 9 and 10, respectively.

	QUANTIFIER	CONDITION	Sentence
<i>wonder</i>	Every	False	Every chef wonders whether or not she will be rated three stars.
		Target	Every chef wonders whether she will be rated zero, one, or two stars.
		True	Every chef wonders whether or not she will be rated one star.
	No	False	No chef wonders whether or not she will be rated one star.
		Target	No chef wonders whether she will be rated zero, one, or two stars.
		True	No chef wonders whether or not she will be rated three stars.
	Exactly	False	Exactly two chefs wonder whether or not they will be rated three stars.
		Target	Exactly two chefs wonder whether they will be rated zero, one, or two stars.
		True	Exactly two chefs wonder whether or not they will be rated two stars.
<i>believe</i>	Every	False	Every chef believes that she might be rated three stars.
		Target	Every chef believes that she might be rated zero, one, or two stars.
		True	Every chef believes that she might be rated one star.
	No	False	No chef believes that she might be rated one star.
		Target	No chef believes that she might be rated zero, one, or two stars.
		True	No chef believes that she might be rated three stars.
	Exactly	False	Exactly two chefs believe that they might be rated three stars.
		Target	Exactly two chefs believe that they might be rated zero, one, or two stars.
		True	Exactly two chefs believe that they might be rated two stars.
Fillers		False	Every chef will be rated at least one star.
		True	No chef will be rated three stars.

Table 9 Stimuli for the Chefs context replication of Experiment 2

	QUANTIFIER	CONDITION	Sentence
<i>wonder</i>	Every	False	Every student wonders whether or not Ms. Adams will teach the 12th-grade geography class next year.
		Target	Every student wonders whether Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Every student wonders whether or not Mr. Brown will teach the 12th-grade geography class next year.
	No	False	No student wonders whether or not Mr. Smith will teach the 12th-grade geography class next year.
		Target	No student wonders whether Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	No student wonders whether or not Ms. Jones will teach the 12th-grade geography class next year.
	Exactly	False	Exactly eight students wonder whether or not Ms. Jones will teach the 12th-grade geography class next year.
		Target	Exactly eight students wonder whether Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Exactly eight students wonder whether or not Mr. Smith will teach the 12th-grade geography class next year.
<i>believe</i>	Every	False	Every student believes that Ms. Adams may be teaching the 12th-grade geography class next year.
		Target	Every student believes that Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Every student believes that Mr. Brown may be teaching the 12th-grade geography class next year.
	No	False	No student believes that Mr. Smith may be teaching the 12th-grade geography class next year.
		Target	No student believes that Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	No student believes that Ms. Jones may be teaching the 12th-grade geography class next year.
	Exactly	False	Exactly eight students believe that Ms. Jones may be teaching the 12th-grade geography class next year.
		Target	Exactly eight students believe that Mr. Smith, Ms. Adams, or Mr. Brown will teach the 12th-grade geography class next year.
		True	Exactly eight students believe that Mr. Smith may be teaching the 12th-grade geography class next year.
Fillers		False	The students think that the school will hire a new geography teacher.
		True	The students can take geography in 12th grade.

Table 10 Stimuli for the Geography context replication of Experiment 2

B.1 Design

Once again, the design was nearly identical to Experiment 2, but VERB was made a within-participant factor and two fillers were added for a total of eight items per participant. QUANTIFIER remained a between-participants factor and we introduced a between-participants CONTEXT factor with two levels (Chefs and Geography). We did not use colors this time, since one reviewer expressed some doubts as to their efficiency to help participants perform the task, and they could have introduced some artifacts.

We followed a randomization procedure similar to that described in Appendix A, but for each of the 6 possible sets of items, we recruited 96 participants, each tested on a different order of these items.²²

B.2 Participants

576 participants (96 for each set of items) were recruited on Amazon’s Mechanical Turk and paid 57¢ each. No two participants saw the same set of items in the same order. The data from 16 participants who reported native languages other than English were discarded. We observed 7 retakes this time, which were discarded from the analysis and not paid. In all, we kept the data from 553 unique participants.

B.3 Results

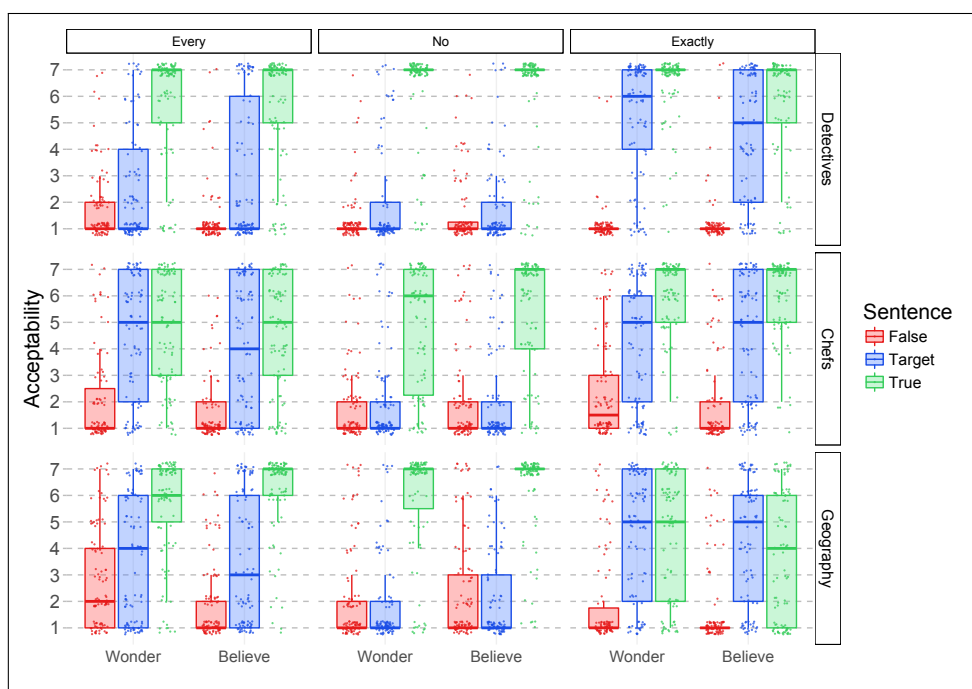


Figure 7 Replication of Experiment 2: Acceptability for each Verb, Quantifier, and Sentence by Context (boxes indicate median and quartiles, each dot represents an individual answer). The results of the original experiment (Detective context) are repeated in the first row.

²² Concretely, we tested all combinations of $\sigma, \sigma', \sigma_1, \sigma_2, \sigma_3$, and only randomly picked one of the two possible σ_5 (see fn 21 for details).

The results are presented in Figure 7. We ran the same analysis as in Experiment 2, but the mixed-effects model included a random slope for VERB since this factor was now within-participant:

(67) Model: `clmm(Answer~(Baseline+Local)*Verb*
Quantifier+(1+Base+Verb|Subject))`

Table 11 presents the results of the model. While the overall pattern for targets didn't differ much from the original experiment, we observed a few differences which are mainly driven by differences in the control items. In particular, the true *every* and *no* controls in the Chefs context and the true *exactly* controls in the Geography context received surprisingly low ratings, which affects our measure of local DIIs in these cases, even though the target sentences behaved very much as in the original experiment.

Nevertheless, we again observed local DIIs under *every* in both contexts, less DIIs under *no* and more DIIs under *exactly*. We observed a small difference in the effect of VERB in the Geography context. Nevertheless, across the three contexts, VERB still didn't reach significance ($\chi^2(9) = 16, p = .066$).

Finally, we fitted a model on data from the quantifier *no* only and found that there was no effect of LOCAL in the Chefs context ($z = -.25, p = .80$), nor in the Geography context ($z = -1.6, p = .10$), unlike in the original Detective context ($z = 2.1, p = .033$).

	β	z-value	p-value
BASELINE	4.46	15.82	< .001***
LOCAL	2.90	12.90	< .001***
VERB	1.49	3.60	< .001***
QUANTIFIER: <i>no</i>	-0.27	-0.90	.370
QUANTIFIER: <i>exactly</i>	-1.33	-3.77	< .001***
CONTEXT: <i>Chefs</i>	0.39	1.34	.179
CONTEXT: <i>Geography</i>	1.29	4.65	< .001***
BASELINE \times VERB	-1.18	-2.17	.030*
LOCAL \times VERB	0.55	1.25	.211
BASELINE \times [QUANTIFIER: <i>no</i>]	1.50	3.67	< .001***
BASELINE \times [QUANTIFIER: <i>exactly</i>]	2.09	4.74	< .001***
LOCAL \times [QUANTIFIER: <i>no</i>]	-2.31	-6.71	< .001***
LOCAL \times [QUANTIFIER: <i>exactly</i>]	2.06	5.32	< .001***
VERB \times [QUANTIFIER: <i>no</i>]	-2.09	-3.52	< .001***
VERB \times [QUANTIFIER: <i>exactly</i>]	-1.85	-2.62	.009**
BASELINE \times [CONTEXT: <i>Chefs</i>]	-1.49	-3.89	< .001***
BASELINE \times [CONTEXT: <i>Geography</i>]	-1.16	-3.09	.002**
LOCAL \times [CONTEXT: <i>Chefs</i>]	-2.52	-8.50	< .001***
LOCAL \times [CONTEXT: <i>Geography</i>]	-0.70	-2.32	.020*
VERB \times [CONTEXT: <i>Chefs</i>]	-0.75	-1.43	.152
VERB \times [CONTEXT: <i>Geography</i>]	-0.34	-0.68	.497
[QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	0.62	1.52	.129
[QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	1.80	3.95	< .001***
[QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	0.11	0.27	.786
[QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	0.15	0.33	.742
BASELINE \times VERB \times [QUANTIFIER: <i>no</i>]	1.51	1.86	.064
BASELINE \times VERB \times [QUANTIFIER: <i>exactly</i>]	2.38	2.71	.007**
LOCAL \times VERB \times [QUANTIFIER: <i>no</i>]	0.03	0.05	.960
LOCAL \times VERB \times [QUANTIFIER: <i>exactly</i>]	0.61	0.79	.430
BASELINE \times VERB \times [CONTEXT: <i>Chefs</i>]	0.71	1.03	.302
BASELINE \times VERB \times [CONTEXT: <i>Geography</i>]	-0.76	-1.12	.264
LOCAL \times VERB \times [CONTEXT: <i>Chefs</i>]	-0.55	-0.94	.346
LOCAL \times VERB \times [CONTEXT: <i>Geography</i>]	-1.28	-2.15	.031*
BASELINE \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	-1.33	-2.38	.017*
BASELINE \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	-1.24	-2.12	.034*
BASELINE \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	-0.53	-0.93	.351
BASELINE \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	-2.68	-4.54	< .001***
LOCAL \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	1.87	4.15	< .001***
LOCAL \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	-0.38	-0.79	.429
LOCAL \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	-0.40	-0.89	.375
LOCAL \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	-1.47	-2.98	.003**
VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	1.18	1.58	.113
VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	1.95	2.32	.020*
VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	0.67	0.93	.353
VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	1.03	1.21	.225
BASELINE \times VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	-1.51	-1.50	.134
BASELINE \times VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	-3.02	-2.82	.005**
BASELINE \times VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	-0.12	-0.11	.910
BASELINE \times VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	-0.03	-0.03	.974
LOCAL \times VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Chefs</i>]	0.43	0.48	.633
LOCAL \times VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Chefs</i>]	-1.56	-1.64	.102
LOCAL \times VERB \times [QUANTIFIER: <i>no</i>] \times [CONTEXT: <i>Geography</i>]	0.64	0.71	.475
LOCAL \times VERB \times [QUANTIFIER: <i>exactly</i>] \times [CONTEXT: <i>Geography</i>]	-0.01	-0.01	.993

Table 11 Results of the model comparing the replications of Experiment 2 to the original experiment. Effects of interest in the original experiment are highlighted in blue, and their interaction with CONTEXT in brown. The standard deviation of the random effects was 1.14 for the intercept, 1.58 for BASELINE, .14 for VERB, and .38 for the interaction.

EARLY ACCESS

Alexandre Cremers
Science Park 107
1098 XG Amsterdam
Netherlands
alexandre.cremers@gmail.com

Floris Roelofsen
Science Park 107
1098 XG Amsterdam
Netherlands
floris.roelofsen@gmail.com

Wataru Uegaki
Van Wijkplaats 4
2311 BX Leiden
Netherlands
w.uegaki@hum.leidenuniv.nl